Encyclopedia of
EDUCATIONAL
PSYCHOLOGY

EDITOR
Neil J. Salkind

VOLUMES 1&2

Encyclopedia of

EDUCATIONAL

PSYCHOLOGY

Encyclopedia of
EDUCATIONAL
PSYCHOLOGY

EDITOR
Neil J. Salkind
UNIVERSITY OF KANSAS

MANAGING EDITOR
Kristin Rasmussen
UNIVERSITY OF KANSAS

Contents

List of Entries

Reader's Guide

This list is provided to assist readers in locating entries on related topics. Some entry titles appear in more than one category.

Intelligence and Intellectual Development

Language Development

Learning and Memory

Organizations

Peers and Peer Influences

Public Policy

Theory

About the Editor

Neil J. Salkind received his Ph.D. from the University of Maryland in Human Development and has been a faculty member at the University of Kansas for 34 years. He has given more than 125 presentations and is the author of several college-level textbooks, including *Statistics for People Who Think They Hate Statistics* (Sage), *Exploring Research and Theories of Human Development* (Sage), and *Exploring Research* (Prentice Hall). He was editor of *Child Development Abstracts and Bibliography* for 13 years and lives in Lawrence, Kansas.

Contributors

Jane L. Abraham
Virginia Tech

Taylor W. Acee
University of Texas at Austin

Gerald R. Adams
University of Guelph

Tufan Adiguzel
Texas A&M University

Lin-Miao L. Agler
University of Southern Mississippi, Gulf Coast

William T. Akers
Aurora University

Michael P. Alfano
University of Connecticut

Benjamin Allen
National Head Start Association

Shannon Altenhofen
Colorado State University

Virginia Johnson Anderson
Towson University

Gabriel Aquino
Skidmore College

Jason Arndt
Middlebury College

Leah M. Rouse Arndt
University of Wisconsin–Milwaukee

Tiffany Lynnette Arrington
University of Kansas

Elise M. Arruda
Claremont Graduate University

Christine Ashby
Syracuse University

David J. Atencio
University of New Mexico

Margarita Azmitia
University of California, Santa Cruz

Jennifer Baggerly
University of South Florida

Melinda E. Baham
Arizona State University

William M. Bart
University of Minnesota

Anne S. Beauchamp
University of Kansas

Robinder P. Bedi
University of Victoria

Ronald A. Beghetto
University of Oregon

Janine Bempechat
Wheelock College

Joyce Benenson
Emmanuel College

Aaron S. Benjamin
University of Illinois at Urbana–Champaign

Eric R. Benson
University of Kansas

Nick Benson
Florida International University

Margit I. Berman
University of Minnesota

Tiffany Berry
Claremont Graduate University

Maximilian B. Bibok
Simon Fraser University

Douglas Biklen
Syracuse University

Zeynep Biringen
Colorado State University

Clancy Blair
Pennsylvania State University

Tracie L. Blumentritt
University of Wisconsin–La Crosse

Sara Bolt
Michigan State University

Trevor G. Bond
Hong Kong Institute of Education

Gary Borich
University of Texas at Austin

Gary Boyd
Concordia University

Jean A. Boyer
Temple University

Jeffery P. Braden
North Carolina State University

Laurie Brady
University of Technology, Sydney

Sanford L. Braver
Arizona State University

Natalie H. Brito
College of William & Mary

Ralph G. Brockett
University of Tennessee

Aaron M. Brower
University of Wisconsin–Madison

Carole W. Brown
Catholic University of America

Scott W. Brown
University of Connecticut

Penny L. Burge
Virginia Tech

Joshua A. Burk
College of William & Mary

James P. Byrnes
Temple University

Judy Cameron
University of Alberta

Miguel Ángel Cano
Texas A&M University

Noel A. Card
University of Arizona

Saul Carliner
Concordia University

Carolyn L. Carlson
University of Kansas

John S. Carlson
Michigan State University

Jeremy I. M. Carpendale
Simon Fraser University

Linda Bryant Caviness
La Sierra University

Tam Chandler
Fielding Graduate University

Carol A. Chapelle
Iowa State University

Zhe Chen
University of California, Davis

Theodore James Christ
University of Minnesota

Tracy Thorndike Christ
Western Washington University

Cathleen Clerkin
University of California, Berkeley

Jill S. M. Coleman
Ball State University

Sarah C. Conklin
Northern Illinois University

Benjamin A. Converse
University of Chicago

Kelly J. Copeland
Missouri State University

Susan R. Copeland
University of New Mexico

Jacqueline Cossentino
University of Maryland

Nelson Cowan
University of Missouri

Lisa W. Coyne
Suffolk University

Jean B. Crockett
University of Florida

David Yun Dai
*University at Albany,
State University of New York*

Matthew J. Davis
Texas A&M University

Brooke Davy
Colorado State University

Edward L. Deci
University of Rochester

Ester Johanna de Jong
University of Florida

Gypsy M. Denzine
Northern Arizona University

Tracey M. Derwing
University of Alberta

Martin John Doherty
University of Stirling

John W. Donahoe
University of Massachusetts

Michelle A. Drefs
University of Calgary

R. F. Drewett
Durham University

Carlton T. Duff
University of Victoria

Martha Dunkelberger
University of Houston

Donna Dunning
University of Victoria

Tina D. Du Rocher Schudlich
Western Washington University

John Warren Eagle
University of Kansas

J. Mark Eddy
Oregon Social Learning Center

Andrew J. Elliot
University of Rochester

Charles H. Elliott
Fielding Graduate University

Nicholas Epley
University of Chicago

Norma D. Feshbach
University of California, Los Angeles

Kimberly Filer
Virginia Tech

Susan Finley
Washington State University

Catherine A. Fiorello
Temple University

Jack M. Fletcher
University of Houston

Jocelyn R. Folk
Kent State University

Sean Alan Forbes
Auburn University

Mary E. Fowles
Educational Testing Service

Jennifer A. Fredricks
Connecticut College

Nanette S. Fritschmann
Lehigh University

Perry N. Fuchs
University of Texas at Arlington

Gilles E. Gignac
University of Western Australia

Jennie K. Gill
University of Victoria

Lisa M. Given
University of Alberta

Mead Goedert
Wayne State University

Ernest T. Goetz
Texas A&M University

Jorge E. Gonzalez
Texas A&M University

Joel B. Goodin
Florida State University

Mileidis Gort
University of Miami

Adele Eskeles Gottfried
California State University, Northridge

Taylor Grant
Colorado State University

Scott L. Graves, Jr.
Bowling Green State University

M. E. Gredler
University of South Carolina

Donald P. Green
Yale University

Vanessa A. Green
University of Tasmania

Michael Guilbault
Howard University

Tracy Hack
Butler County Community College

Paul John Hager
University of Technology, Sydney

Laura S. Hamilton
RAND Corporation

Frederick D. Harper
Howard University

Vincent J. Harper
University of Oklahoma

Paula Hartman
University of Texas at Austin

S. Alexander Haslam
University of Exeter

Billy Hawkins
University of Georgia

Felicity Ann Haynes
University of Western Australia

Elizabeth V. Heath
Allegheny County Public Schools

Larisa Heiphetz
Pennsylvania State University

Jovan Hernandez
University of Iowa

Deneil Hill
Colorado State University

Edward Raymond Hirt
Indiana University

Renee Hobbs
Temple University

Susan D. Holloway
University of California, Berkeley

Donald Homa
Arizona State University

Saa Hoon Hong
University of Minnesota

Ryan Honomichl
University of California, Davis

Richard E. Hult
University of South Carolina

Matthew Huss
Creighton University

Royce A. Hutson
Wayne State University

Doyna Illmer
Catholic University of America

Karen Kurotsuchi Inkelas
University of Maryland

Lori Jackson
Georgia State University

Danielle Johnson
University of Kansas

Kent Johnson
Morningside Academy

Paula Johnson
Western Washington University

Tricia S. Jones
Temple University

Diana Joyce
University of Florida

Debra M. Hernandez Jozefowicz-Simbeni
Wayne State University

Charles Kalish
University of Wisconsin–Madison

Michael L. Kamil
Stanford University

James S. Kaminsky
Auburn University

Kentaro Kato
University of Minnesota

Alan E. Kazdin
Yale University

Karl Kelley
North Central College

Sean Kelly
University of Notre Dame

John J. Ketterer
Jacksonville State University

Christopher B. Keys
DePaul University

Edward "Skip" Kifer
University of Louisville

Kerri Lynn Kim
University of Kansas

Neal Kingston
University of Kansas

David M. Klieger
University of Minnesota

Barbara Korth
Indiana University

S. Kathleen Krach
University of Nevada, Las Vegas

Michael W. Kramer
University of Missouri

Marie Kraska
Auburn University

Bill Kring
Educational Service District 112

Rachael Elizabeth Kroening
Missouri State University

Nathan R. Kuncel
University of Minnesota

Paul G. LaCava
University of Kansas

Jeffrey A. Lackney
School Design Research Studio, University of Wisconsin–Madison

Martha J. Larkin
University of West Georgia

Philip J. Lazarus
Florida International University

Jorja Leap
University of California, Los Angeles

Jacqueline P. Leighton
University of Alberta

Rosalind Levačić
University of London

Chantal Levesque
Missouri State University

Michael K. Lindell
Texas A&M University

Eric W. Lindsey
Pennsylvania State University, Berks

Louis G. Lippman
Western Washington University

William Ming Liu
University of Iowa

Lailei Lou
University at Albany,
State University of New York

Patricia A. Lowe
University of Kansas

Ruth Luckasson
University of New Mexico

Ronald A. Madle
Pennsylvania State University

Joseph P. Magliano
Northern Illinois University

John Malouff
University of New England

James E. Martin
University of Oklahoma

Joan M. Martin
University of Victoria

Patricia G. Mathes
Southern Methodist University

Matthew Jared Mayer
Rutgers University

Michael E. McCarty
Texas Tech University

Katherine E. McDonald
Portland State University

Adriana G. McEachern
Florida International University

Kara E. McGoey
Duquesne University

Robyn McKay
University of Kansas

John E. McKenna
St. Francis Xavier University

Robert F. McMorris
University at Albany,
State University of New York

Guy R. McPherson
University of Arizona

Frances K. McSweeney
Washington State University

Michael H. McVey
Eastern Michigan University

Chris S. Meiers
University of Kansas

Noorfarah Merali
University of Alberta

Peter Merrotsy
University of New England

Darcy Miller
Washington State University

Jeremy K. Miller
Willamette University

Joyce E. Miller
Texas A&M University

M. David Miller
University of Florida

Claudia L. Moreno
Rutgers University

Gale M. Morrison
University of California, Santa Barbara

Margaret Moustafa
California State University, Los Angeles

Ulrich Müller
University of Victoria

Karen D. Multon
University of Kansas

Eric S. Murphy
University of Alaska, Anchorage

Angela K. Murray
University of Kansas

Shannon Myrick
Portland State University

Stacey Neuharth-Pritchett
University of Georgia

Christopher R. Niileksela
University of Kansas

David P. O'Brien
*Baruch College and the Graduate Center,
City University of New York*

Nora V. Odendahl
Educational Testing Service

Jessica Oeth
University of Kansas

Corrie Orthober
Bellarmine University

Lora E. Park
*University at Buffalo,
State University of New York*

Carol S. Parke
Duquesne University

Kay T. Payne
Howard University

Nanette L. Perrin
*Community Living Opportunities,
    Early Child Autism Program*

Paul J. Perry
Northern Illinois University

David P. Peterson
3M Company

Gail B. Peterson
University of Minnesota

Michael Petkovich
Independent Consultant

Andy V. Pham
Michigan State University

W. David Pierce
University of Alberta

Jonathan A. Plucker
Indiana University

Richard S. Prawat
Michigan State University

Olga Acosta Price
George Washington University

Shirley Mary Pyon
University of North Carolina at Chapel Hill

Patricia D. Quijada
University of Texas, San Antonio

Audrey M. Quinlan
Seton Hill University

Jennifer M. Raad
University of Kansas

Manuel Ramirez, III
University of Texas at Austin

Matthew Soldner
University of Maryland

Rachel Milstein Sondheimer
United States Military Academy

Samuel Y. Song
University of North Carolina at Chapel Hill

Caile E. Spear
Boise State University

Jason M. Stephens
University of Connecticut

Robert J. Sternberg
Tufts University

Robert J. Stevens
Pennsylvania State University

William T. Stokes
Lesley University

Heather Stopp
Claremont Graduate University

Elizabeth M. Street
Central Washington University

Hoi K. Suen
Pennsylvania State University

Daniel Suitor
University of Kansas

Xiaoyuan Tan
*University at Albany,
State University of New York*

Sophia Tani-Prado
Texas A&M University

Rebecca J. Thompson
Temple University

Robert M. Thorndike
Western Washington University

Mary Ellen Tillotson
Rhode Island College

Keith Topping
University of Dundee

Jeannine E. Turner
Florida State University

Amanda M. VanDerHeyden
University of California, Santa Barbara

Peggy N. Van Meter
Pennsylvania State University

Kimberly J. Vannest
Texas A&M University

Theresa K. Vescio
Pennsylvania State University

Peter M. Vishton
College of William & Mary

Nagalapura S. Viswanath
West Texas A&M University

John Walsh
University of Victoria

Noreen M. Webb
University of California, Los Angeles

Claire Ellen Weinstein
University of Texas at Austin

Stanley Jerome Weiss
American University

Greg W. Welch
University of Kansas

Introduction

Educational psychology is a special field of endeavor because it strives to apply what is known about many different disciplines to the broad process of education. In the most general terms, you can expect to find topics in this area that fall into the categories of human learning and development (across the life span), motivation, measurement and statistics, and curriculum and teaching. More specifically, the educational psychologist studies such topics as aggression, the relationship between poverty and achievement in schools, lifelong learning, quantitative methods, and emerging adulthood. Educational psychology is truly a diverse and fascinating field of study and unlike other social and behavioral sciences. Its significance for application to the real needs of both children and adults cannot be overestimated.

The importance of all these topics is not limited to the college classroom or academic lecture circuit. Rather, the ability to understand complex issues such as vouchers, early intervention, inclusion, cultural diversity, and the role of athletics in the schools (to mention only a few examples) carries important implications for public policy decisions. The encyclopedia you have in your hands includes some technical topics related to educational psychology, but for the most part, it focuses on those topics that evoke the interest of the everyday reader.

Although there are hundreds of books about different topics in education and educational psychology and there are thousands of university and private researchers pursuing more information about these topics, most of the available information tends to be found in scholarly books and scholarly journal articles—usually out of the reach of the everyday person. In fact, there are few comprehensive overviews of the field of educational psychology, and the purpose of this multivolume *Encyclopedia of Educational Psychology* is to share this information in a way that is, above all, informative without being overly technical or intimidating.

Through more than 275 contributions, experts provide overviews and explanations of the major topics in the field of educational psychology.

How were these topics selected to be included in this encyclopedia? The underlying rationale for topic selection and presentation comes from the need to share subjects that are rich, diverse, and deserving of closer inspection with an educated reader who may be uninformed about educational psychology. Within these pages, the contributors and I provide the overview and the detail that we feel is necessary to become well acquainted with topics that fairly represent the entire field.

Like many encyclopedias, the *Encyclopedia of Educational Psychology* is organized in alphabetical order, from A through Z. However, there are particular themes under which the information and the entries could be organized conceptually. These themes, or major topic areas, constitute the Reader's Guide, which appears on page xi. Categories such as Classroom Management; Ethnicity, Race, and Culture; Families; Intelligence and Intellectual Development; Learning and Memory; and Peers and Peer Influences are only a few that help organize the entire set of contributions.

The Process

Sage developmental editors Diana Axelsen and Carole Maurer, the encyclopedia's managing editor Kristin Rasmussen, and I started the process of compiling all the entries in the encyclopedia by asking experts in the general field, as well as the more specific areas (again, see the Reader's Guide), what topics they feel should be included to make a reference work that provides a general overview of the field. We tried to ensure that these entries included topics that would be of interest

to a general readership and not terms that were too highly technical or too far removed from the interests of the everyday reader. This list was reviewed and revised many times, until we felt that it was a comprehensive set of topics that best fit the vision for the encyclopedia.

It was no surprise that this list would be edited and revised as we worked and as authors were recruited to write particular entries. Enthusiastic authors would suggest adding topics that might have been overlooked as well as removing topics that might have no appeal. These suggestions were taken into consideration as the final list that forms the A through Z collection was assembled.

The next step was to assign a length to each entry, which ranges from 1,000 words for shorter articles (such as the one on Parent–Teacher Conferences) to more than 5,000 words for major entries (such as the one on Literacy). The scope of different entries varies depending on their importance to the field and the amount of information we thought important to include. In between, there are entries that are 2,000 and 3,000 words in length. Sometimes authors would request that the length be extended because they had so much information they wanted to include and they felt that the limitation on space was unwarranted. In most cases, it was not a problem to allow such an extension.

The final step was the identification of authors. This took place through a variety of mechanisms, including experts' recommendations, professional and personal experiences of both the managing editor and me, contact with authors of journal articles and books that focus on a particular area directly related to the entry, and requests for referrals from other individuals who are well known in the field. It is a testament to how deeply these authors are committed to their own work and to sharing it with others that they gave freely of their time and advice.

Once authors were identified and invited, and once they confirmed that they could participate, they were sent detailed instructions and given a deadline for the submission of their entry. The results, after editing, layout, and other production steps, are in your hands.

How to Use This Book

Like a good meal, a book is meant to be fully enjoyed, and whereas most people believe that encyclopedias are only used for reference purposes, we hope that this two-volume encyclopedia is an easy one to sit with and browse.

A primary goal of creating this set of volumes was to open up the broad discipline of educational psychology to a wide and general audience. That's why you will find topics that are of particular interest to the general public, such as vouchers, Head Start, divorce, learning communities, and charter schools.

Take these books, find a comfortable seat in the library, browse through the topics, and read the ones that catch your eye. We're confident that you'll continue reading about the field and will use the suggestions for such contained in the Further Readings list following each entry.

Should you want to find a topic within a particular area, consult the Reader's Guide, which organizes entries within this two-volume set into general categories. Using this tool, you can quickly move to an area or a specific topic that you find valuable and of interest.

Acknowledgments

My most appreciative thanks to Diane Axelsen and Carole Maurer. Their title is developmental editor, but they are the engineers, timetable keepers, track fixers, customs agents, and chefs who keep the train running and the people who are on it (the editorial staff and all the contributors) happy and productive. It is not an exaggeration to say that this encyclopedia would not have happened without their selflessness, diligence, patience (such as, "Okay Carole, I promise I'll get it to you by tomorrow"), and nurturance. I am in their debt.

My sincere thanks and appreciation go to the managing editor, Kristin Rasmussen, an advanced graduate student at the University of Kansas, who managed the submissions from the recruitment of authors to the transmission of the final documents to Sage. Kristin completed these tasks with enthusiasm, initiative, and perseverance in answering endless questions through thousands of e-mails to hundreds of authors. Thank you sincerely.

Neil J. Salkind
University of Kansas
June, 2007

A

When I was a boy of fourteen, my father was so ignorant I could hardly stand to have the old man around. But when I got to be twenty-one, I was astonished at how much he had learned in seven years.

—Mark Twain, "Old Times on the Mississippi," *Atlantic Monthly*, 1874

ABSTINENCE EDUCATION

Abstinence education advocates abstinence as the 100% sure way to prevent pregnancy and the transmission of sexually transmitted diseases (STDs). The U.S. teen pregnancy rates have been decreasing since the 1990s, but of all of the developed countries in the world, the United States still has the highest teen pregnancy rates. The Centers for Disease Control and Prevention estimate that approximately 19 million new STDs occur each year, almost half of them among young people ages 15 to 24. Teens with STDs are impacted emotionally and physically, and the annual medical cost directly attributed to STDs in the United States is well over $13 billion. Communities are looking for answers on how to address the complex issues of teen pregnancy and STDs. Abstinence education is seen as one such answer.

The premise of abstinence education is that abstinence is the best choice for youth when it comes to making sexual decisions. Schools and communities can decide what type of programming to provide to their young people. The current federal initiatives are driving the abstinence-until-marriage initiatives, and anyone receiving federal funding must adhere to specific guidelines. The evaluation of these programs shows some short-term impacts on attitudes and behavioral intent, but long-term studies on behavior are mixed. This entry provides a general overview of sexuality education, abstinence education, federal funding for abstinence programming, guidelines for selecting programs, and evaluation of abstinence programs.

Sexuality and Abstinence Education

Parents are the primary educators of their children. Some parents may not feel comfortable broaching the topics with their children, may not have the factual knowledge to share with their children, or may not know how to talk to their children in a developmentally appropriate manner. Who else can teach the children? The vast majority of children attend public school, and schools are a logical place to provide sexuality education. Programming may be taught by teachers, health educators, nurses, doctors, or other credentialed professionals. No matter who is teaching the children, it is important for the school to verify the accuracy of the curriculum, evaluate the credentials of the provider, and be sure the curriculum meets with school district policies.

Sexuality education guidelines are typically established at the state level and carried out at the local level. Almost every state mandates sexuality education, and some mandate abstinence education. Districts and communities may choose from a range of

programs, from comprehensive sexuality education programs—which may cover birth control, a range of sexual behaviors, gender identity, life skills, and anatomy and physiology—to abstinence-only-until-marriage programs, which only discuss sexual activity within the context of marriage. School districts typically create policies delineating what type of sexuality education will be taught in what grades and what topics will be covered. Most school districts allow students, with parental permission, to opt out of sexuality education.

Sexuality education has changed from the broader ranging comprehensive sexuality education in the 1970s to the current, more narrowly focused, federally funded abstinence-only-until-marriage programs of the late 1990s and today. The change in program focus can be attributed to a concern among some parents and communities that comprehensive sexuality education was teaching students how to have sex or was sending the mixed message "Do not have sex—but if you do, be sure to use protection." Some of the disfavor arose from not clearly delineating what could and should be taught in the schools, that is, concerns over what was developmentally appropriate and what was the purview of the parent. Some people felt abstinence education was a means of establishing moral purity and strengthening marriage.

Abstinence education proponents support the one message of abstinence from sex and do not want students to receive the mixed message "Yes, remain abstinent—but if you are not abstinent, then remember to use birth control." Many abstinence education programs do not discuss birth control except to describe their failure rates. Components of abstinence education programs can vary, but generally the focus is on the harm that comes from sexual intercourse and early sexual involvement. Depending on the school district, curricula components vary and may or may not include anatomy and physiology, life skills, communication skills, or refusal skills components.

The curricula are created by a variety of people, including teachers, school districts, health educators, faith-based groups, private businesses, or community members.

Abstinence-plus proponents believe abstinence is the best choice for young people when it comes to making decisions about sex. The "plus" in abstinence-plus often includes life skills components such as goal setting, life planning, communication, anatomy and physiology, and information on contraception. The plus component refers to the ability of the teacher to answer student questions, provide information about contraceptives, or possibly refer the student for information on contraceptives.

The philosophy of the current abstinence-only-until-marriage, also known as abstinence-only, programs is that sexual intercourse and sexual activity should happen only between a man and woman and only when they are married. There is no discussion of contraception except to discuss failure rates, and no programs may advocate for the use of contraceptives. Each program adheres to a set of guidelines as established by law in 1996. The next section discusses federal funding of abstinence-only programs and the program guidelines.

Federal Funding

The first federal funding of abstinence programs was created in 1981 with the Adolescent Family Life Act as Title XX of the Public Health Service Act. According to the Office of Adolescent Pregnancy Programs, the Adolescent Family Life (AFL) program supports demonstration projects to develop, implement, and evaluate program interventions to promote abstinence from sexual activity among adolescents and to provide comprehensive health care, education, and social services to pregnant and parenting adolescents. The program supports two basic types of demonstration projects: (1) prevention demonstration projects to develop, test, and use curricula that provide education and activities designed to encourage adolescents to postpone sexual activity until marriage; and (2) care demonstration projects to develop interventions with pregnant and parenting teens, their infants, male partners, and family members in an effort to ameliorate the effects of too-early-childbearing for teen parents, their babies, and their families. The AFL program also funds grants to support research on the causes and consequences of adolescent premarital sexual relations, pregnancy, and parenting. The Title XX funds not only help the teens and families they serve directly, but they also provide valuable information and evaluation findings that can serve as a basis for future strategies. Every program that receives AFL grant funds is required to include an independent evaluation component. This ensures that the lessons learned by each community will benefit others in the future.

In 1996, federal abstinence education programs narrowed the definitions of abstinence when Section 510(b) of Title V of the Social Security Act,

P.L. 104–193 was signed into law by President Bill Clinton. The abstinence-only federal funding was created as part of "welfare reform," or the Temporary Assistance for Needy Families Act (TANF). This third funding stream provides grants to states for abstinence-only-until-marriage programs. The abstinence-only-until-marriage educational or motivational programs must adhere to the following eight criteria as established by law:

1. Has as its exclusive purpose teaching the social, psychological, and health gains to be realized by abstaining from sexual activity

2. Teaches abstinence from sexual activity outside marriage as the expected standard for all school-age children

3. Teaches that abstinence from sexual activity is the only certain way to avoid out-of-wedlock pregnancy, sexually transmitted diseases, and other associated health problems

4. Teaches that a mutually faithful monogamous relationship in the context of marriage is the expected standard of sexual activity

5. Teaches that sexual activity outside of the context of marriage is likely to have harmful psychological and physical effects

6. Teaches that bearing children out-of-wedlock is likely to have harmful consequences for the child, the child's parents, and society

7. Teaches young people how to reject sexual advances and how alcohol and drug use increases vulnerability to sexual advances

8. Teaches the importance of attaining self-sufficiency before engaging in sexual activity

As of 1997, all abstinence projects funded under the Office of Adolescent Pregnancy Prevention must adhere to the eight criteria. The funded projects must be evaluated, and the curricula must be medically accurate.

Title V grantees cannot provide educational programming that goes against any one of the criteria listed previously in this section, but the states have the latitude to focus on only a few of the criteria. The states can direct the funding to schools, community-based organizations, health districts, media campaigns, or faith-based entities. Each state has the discretion to decide who receives the funding, how programs are delivered, and if and how they will be evaluated.

In October 2000, the federal government expanded the abstinence-only projects and created Special Projects of Regional and National Significance–Community-Based Abstinence Education (SPRANS–CBAE). The SPRANS grants are awarded to states and community organizations and can fund only abstinence-only-until-marriage programs. In 2005 oversight of the SPRANS–CBAE grants was moved from the Maternal and Child Health Bureau to the Administration for Children and Families (ACF); both are within the U.S. Department of Health and Human Services (HHS). This program is now known as Community-Based Abstinence Education (CBAE). The program requirements have been tightened; to receive funding, each program now must adhere to all eight criteria. Process evaluation is now required for new grantees, but programs do not have to measure impact on program participants. So how does one select an abstinence program?

Selecting and Evaluating Programs

Abstinence education programs are evolving, and the types of programs offered are vast in terms of quality, cost, and effectiveness. When choosing a program, educators can look for teacher training, medical accuracy, evidence of effectiveness with similar populations, costs, whether the program is theoretically based, and how it fits with community guidelines. Other considerations include whether the program is taught by a teacher or peers and how much time can be devoted to the program. Programs can range from a single 1-hour presentation to 25 sessions over 5 weeks. The curricula and desired learning outcomes should coincide. A final component is the effectiveness of the program; currently, long-term research on the effectiveness of abstinence education is limited to a few programs, and the results are mixed. Few rigorous, long-term studies focusing on behavioral outcomes have been conducted on specific abstinence-only programs. The Office of Adolescent Pregnancy Programs (OAPP) does require all of its abstinence programs to be evaluated, and most of the published abstinence evaluations are former OAPP projects. Most of the programs have shown immediate differences in attitudes and intent to remain abstinent. One problem with the evaluations of many abstinence programs is the limited ability to assign students to experimental or control groups, the presence of follow-up evaluations, and the limited sample sizes.

Research is needed to determine the long-term impacts on behavior.

Caile E. Spear

See also Evaluation; Gender Identity; Sex Education

Further Readings

Administration for Children and Families, U.S. Department of Health and Human Services. (2006). *ACF grant opportunities.* Retrieved March 30, 2007, from http://www.acf.hhs.gov/grants/index.html

Centers for Disease Control and Prevention. (2005). *Trends in reportable sexually transmitted diseases in the United States.* Retrieved April 1, 2007, from http://www.cdc.gov/std/stats/trends2005.htm

Chesson, H. M., Blandford, J. M., Gift, T. L., Tao, G., & Irwin, K. L. (2004, March). *The estimated direct medical cost of STDs among American youth, 2000.* Paper presented at the 2004 National STD Prevention Conference, Philadelphia.

Goodson, P., Pruitt, B. E., Suther, S., Wilson, K., & Buhi, E. (2006). Is abstinence education theory based? *Health Education and Behavior*, *33*, 252–271.

National Campaign to Reduce Teen Pregnancy. (2007). *United States birth rates for teens, 15–19, 1991–2005.* Retrieved April 1, 2007, from http://www.teenpregnancy.org/resources/data/brates.asp

Office of Population Affairs, Office of Adolescent Pregnancy Programs. (n.d.). *OPA legislation.* Retrieved April 1, 2007, from http://opa.osophs.dhhs.gov/legislation.html

Santelli, J. S., Duberstein Lindberg L., Finer, L., & Singh, S. (2007). Explaining recent declines in adolescent pregnancy prevention in the United States: The contribution of abstinence and improved contraceptive use. *American Journal of Public Health*, *97*, 150–156.

Young, M., & Penhollow, T. (2006). The impact of abstinence education: What does the research say? *American Journal of Health Education*, *37*, 194–202.

ACCELERATION

The classical understanding of the term *acceleration* is progress through an educational program at a rate faster, or at an age younger, than conventional. This is now referred to, more appropriately, as *academic acceleration* and is based on the premise that each child has a right to realize his or her potential.

Academic acceleration is valid pedagogy, is grounded in and supported by research, and is an appropriate response to the educational and social needs of a student whose cognitive ability and academic achievement are several years beyond those of their age-peers. Yet worldwide it is an educational option little used. Even though the research on acceleration is so uniformly and distinctly positive and the benefits of well-administered acceleration are so unequivocal, educators are reluctant to accelerate children, and some educational systems proscribe its transparent use.

This entry presents an outline of current theory of academic acceleration through a discussion of a curriculum for gifted students, the benefits of acceleration, a model for acceleration, guidelines for implementing an acceleration program, and ongoing issues related to the practice of acceleration.

Curriculum for Gifted Students

The literature is adamant: Gifted students are exceptional students who have three basic educational needs. They require the provision of a curriculum that is substantially and qualitatively differentiated; that is prescribed, planned, articulated, permanent, ongoing, and defensible; that is based on students' exceptionality; and that is predicated on the needs of each student. Gifted students require accelerated, enriched, and challenging learning experiences, with carefully planned, relevant enrichment and with content acceleration to the level of each student's ability. They also require counseling and guidance to foster cognitive and affective growth. Whereas most teachers and researchers involved with the education of gifted students agree that gifted students do require a differentiated curriculum, there is passionate debate concerning the form that this provision should take.

Counseling certainly is important for the social and emotional development of the gifted student and should be part of the framework for any program devised for gifted students.

Academic enrichment is worthwhile for most students and should not be offered to gifted students only. Relevant academic enrichment requires the provision of a program specifically designed for the individual. For gifted students, this will naturally entail advanced material and higher-level treatment of topics within their area of special aptitude, and the more relevant and excellent the enrichment is, the more it calls for acceleration of subject matter or grade placement later. Indeed, acceleration may well be the most appropriate form of enrichment.

The notion of academic acceleration is evidently contentious, with an enormous hiatus existing between what research has revealed and what most practitioners believe and do. The literature uniformly emphasizes that academic acceleration should form an integral component of a school's program for gifted students, complementing enrichment programs and provisions and following relevant enrichment. Resistance to academic acceleration, especially through concerns for the social and emotional development of the accelerated student, is not grounded in research. Clearly, educators need to be aware of the empirical research on the positive effects of academic acceleration. Moreover, the literature carefully points out that academic acceleration appears to be the best and most feasible method for providing a challenging, rewarding, and ongoing education that matches a gifted student's academic and intellectual ability and comes closest to meeting his or her educational, social, and emotional needs.

Whatever the status of the debate, if service delivery is predicated on a gifted student's precocious development and educational needs, then a differentiated curriculum should be challenging and educationally relevant and should be adapted by acceleration, enrichment, sophistication, and novelty. Such an eclectic approach to programming for the gifted will be employed within an integrative framework, adaptable to the cognitive and affective needs of the individual.

The issue may be placed in perspective by correctly noticing that a gifted student is *already accelerated* and that what is accelerated through academic acceleration is simply the student's progress through the formal school curriculum. The key point is that matching the curriculum to the student's abilities is not acceleration per se, but rather it is a developmentally appropriate teaching practice.

Types of Academic Acceleration

Grade skipping is but one example of academic acceleration. In practice, the range and types of academic acceleration also include early entrance to school, continuous progression, self-paced instruction, correspondence courses, combined classes, multiage classes, curriculum compacting, curriculum telescoping, extracurricular programs, mentorships, content acceleration, subject acceleration, credit by examination, concurrent enrollment, advanced placement (an American practice with few equivalents elsewhere),

early access to advanced-level studies while still at school, and early entrance to university.

Academic acceleration, therefore, refers to any of the ways by which a gifted student engages in the study of new material that is typically taught at a higher grade level than the one in which the child is currently enrolled, covers more material in a shorter time, and accordingly is seen to be *vertical* provision for gifted students. Implicit is the assumption that gifted students, who perform, or reflect the potential to perform, at advanced skill levels should be studying new material at levels commensurate with their levels of ability. Because a common characteristic of gifted students is their ability to learn at a fast rate, acceleration is seen to be a fundamental need of a gifted student and, in some form, should be an integral part of every gifted program.

A model for academic acceleration may refer to service delivery, whereby a standard curriculum experience is offered to a gifted student at a younger age or earlier grade than usual. Or it may refer to curriculum delivery, which involves increasing the pace of presentation of material, either in the regular classroom or in special classes. In either case, programs for academic acceleration allow the examination of content in greater depth, give access to subject matter at levels of greater conceptual difficulty, and should provide instruction that individually and explicitly matches the achievement levels, ability, interests, and learning style of the gifted student.

Benefits of Acceleration

Academic acceleration has several administrative benefits. It is a readily available and inexpensive educational option. It is a way of giving recognition for a student's advanced abilities and accomplishments. It increases learning efficiency, learning effectiveness, and productivity; it gives a student more choice for academic exploration; and it may give increased time for a career.

The report *A Nation Deceived* presents an excellent summary of recent research supporting the academic and affective benefits of well-administered acceleration. From this research come four important findings that are strong and clear and unequivocal.

First, acceleration is consistently and highly effective for academic achievement. No studies have shown that enrichment programs or provisions give more benefits to gifted students than methods of

acceleration. Academic benefits do arise from ability grouping accompanied by a differentiated curriculum, but the greatest benefit comes from academic acceleration. That is, accelerated gifted students, regardless of which form of acceleration is used, significantly outperform students of similar intellectual ability who have not been accelerated.

Second, there is no research to support the claim of maladjustment from acceleration. Despite the preponderance of evidence in favor of academic acceleration, concern about the social and emotional adjustment of accelerated students persists. This concern is cited by both teachers and administrators as the primary reason for opposition to academic acceleration. However, research finds no evidence to support the notion that social and emotional problems arise through well-run and carefully monitored acceleration programs.

Third, acceleration is usually effective in terms of affective adjustment. For many students, it removes them from difficult social situations and from unchallenging and inappropriate educational contexts. It exposes the student to a new peer group and, in fact, significantly increases the chances of a gifted student forming close and productive social relationships with other students. That is, academic acceleration goes a long way to meeting the social and emotional needs of the gifted student who uses it.

Fourth, a gifted student who is not accelerated when it is appropriate may well experience educational frustration and boredom; have reduced motivation to learn; develop poor study habits; have lower academic expectations, achievement, and productivity; express apathy toward formal schooling; drop out prematurely (there is at least some anecdotal evidence to support this); and/or find it difficult to adjust to peers who do not share advanced interests and concerns. That is, rather than expressing concern over potential socioemotional maladjustment arising from acceleration, teachers and administrators need to be concerned about the probability of maladjustment effects resulting from inadequate intellectual challenge.

A Model for Academic Acceleration

It is important to see academic acceleration not as a single intervention but rather as an ongoing, holistic, whole-school process necessarily involving the student, the student's caregivers, and his or her teachers.

Accordingly, Peter Merrotsy has developed an acceleration model that recommends six steps toward a better curriculum for gifted students: identification, communication, a negotiated curriculum, academic acceleration, access to advanced courses while still at school, and support.

Identification

Identification of a gifted student should imply that educational action will take place. It needs to be remembered that identification is notoriously unreliable, especially for gifted students from a background of disadvantage (e.g., low socioeconomic status or forced cultural minority status). That is one of the reasons why it is important to have a broad, inclusive curriculum and to have students involved in making decisions about their curriculum.

Communication

Each gifted student has a right to know the curriculum options and pathways available to him or her. Information about enrichment programs, extracurricular activities, meeting outcomes in alternative ways, high-level courses, senior courses, academic acceleration, and access to advanced-level courses while still at school should be clearly communicated to gifted students, and indeed to all students and their caregivers.

A Negotiated Curriculum

Gifted students are in a position to make informed decisions about their education. They should be actively involved in decision-making processes concerning their curriculum. Gifted students should be empowered to negotiate their curriculum.

Academic Acceleration

Academic acceleration appears to be the best and most feasible method for providing a challenging, rewarding, and continuous education which matches a gifted student's academic ability and comes closest to meeting his or her educational—intellectual, social, and emotional—needs. In order to provide acceleration options it may be necessary to change the organization of the school's curriculum and, in some cases, to change systemic policy.

Advanced-Level Courses

Access to advanced-level courses while still at school is an appropriate and natural progression for a gifted student who has academically accelerated. It is important to remember that an accelerated student could choose instead to study a greater number of secondary subjects, complete fewer secondary subjects but in greater depth, or take a year off, perhaps as an exchange student in another country. Whichever option is pursued, careful long-term planning, clear communication, and a negotiated curriculum are needed.

Support for Gifted Students

To help their intellectual, social, and emotional development, academically accelerated students need appropriate support, in terms of policy, administration, coordination of courses, enriched educational experiences, access to high-level courses, access to specialist teachers, tutors, counselors and mentors, and resources. In particular, gifted students from disadvantaged backgrounds, and from rural and isolated settings, need financial support so that they have access to resources and to educational experiences and opportunities enjoyed by others.

Guidelines for Acceleration

The *Iowa Acceleration Scale* offers a thoughtful and careful objective guide for whole-grade, academically accelerated progression and is supported by research and many repetition studies. There are four critical items: If a student's measured IQ is below 120, if a sibling is either in the same grade from which the student will accelerate or in the new grade to which the student will accelerate, or there is any antipathy by the student, then whole-grade acceleration is not recommended. School history; an assessment of ability, aptitude, and achievement; academic and developmental factors; interpersonal skills; and attitude and support by the school and family are then taken into account in order to give, or not give, as the case may be, a recommendation for whole-grade acceleration. If whole-grade acceleration is not recommended, then advice is available on the suitability of other forms of acceleration or on enrichment and extension programs.

Ongoing Issues

Two key issues need to be addressed worldwide if gifted students are to gain adequate access to a curriculum that includes options for academic acceleration. These issues need to be addressed to overcome the impact of social and cultural disadvantage and to give equity of access to appropriate educational programs for gifted students.

First, the findings of research concerning the academic and affective benefits of well-administered acceleration programs need to be accepted by educational administrators, communities, and teachers. System-organizational patterns of social grouping and the lockstep method of promotion constitute an effective barrier to the development of giftedness, suggesting the deep and urgent need for more flexible forms of school organization that ensure continuity of experience based on criteria other than age or years of attendance and that permit student progression based on individual development and performance.

Second, the end result or consequence of acceleration must be appropriately supported and managed by the education system. For example, with respect to advanced-level subjects studied while still at school, clarification is needed concerning equity of access, which can only be maintained through flexible forms of delivery and alternative modes of study; recognition that they constitute a formal component of secondary school studies, with continuity and articulation of curriculum; the status of secondary students who have completed advanced level units of studies, inter alia that they are still eligible for university entrance scholarships; and credit transfer.

Peter Merrotsy

See also Attachment; Cognitive Development and School Readiness; Emotional Development; Gifted and Talented Students

Further Readings

Assouline, S., Colangelo, N., Lupkowski-Shoplik, A., Lipscombe, J., & Forstadt, L. (2003). *Iowa Acceleration Scale: Manual: A guide for whole-grade acceleration* (2nd ed.). Scottsdale, AZ: Gifted Psychology Press.

Colangelo, N., Assouline, S., & Gross, M. (Eds.). (2004). *A nation deceived: How schools hold back America's brightest students* (2 vols., The Templeton National Report on Acceleration). Iowa City, IA: Belin Blank International

Center for Gifted Education and Talent Development. Retrieved April 20, 2007, from http://nationdeceived.org

Gross, M. (2004). *Exceptionally gifted children* (2nd ed.). London: RoutledgeFalmer.

Gross, M., & van Vliet, H. (2003). *Radical acceleration of highly gifted children.* Sydney, Australia: University of New South Wales, Gifted Education Research, Resource and Information Centre.

Merrotsy, P. (2002). *Appropriate curriculum for academically accelerated students: Listening to the case studies of gifted students.* Unpublished doctoral thesis, Northern Territory University, Darwin, Australia.

Merrotsy, P. (2003). Acceleration: Two case studies of access to tertiary courses while still at school. *TalentEd, 21*(2), 10–24.

Southern, T., & Jones, E. (Eds.). (1991). *The academic acceleration of gifted children.* New York: Teachers College Press.

ACCULTURATION

Acculturation is a complex process that includes those phenomena that result when groups of individuals having different cultures come into continuous firsthand contact, with subsequent changes in the original cultural patterns of either or both groups. The study of acculturation was originally of interest to the fields of anthropology and sociology, focusing on changes occurring at a group level. However, acculturation incorporates changes at the social, group, and individual levels. Later, other fields such as psychology examined acculturation at an individual level. The concept of individual acculturation is also referred to as *psychological acculturation*, which is explained as a change in attitudes, behaviors, beliefs, and values. In relation to acculturation, scholars have identified six areas of functioning that are directly affected by acculturation: language, cognitive styles, personality, identity, attitudes, and acculturative stress. Although acculturation is usually linked to cross-cultural relocations, as with immigrants and refugees, it may take place in numerous sociocultural contexts among a variety of groups. Acculturation is not restricted to those who embark on geographical movement; it can occur in stationary communities, such as in the case of indigenous or native people and ethnic groups in pluralistic societies.

Dimension of Acculturation

Whereas many scholars and research studies have focused on acculturation at a behavioral dimension, it is important to recognize that acculturation can affect other areas as well. Spoken language preference, television program preference, and participation in cultural activities are all ways in which acculturation can be experienced at a behavioral level. Additionally, acculturation can be experienced at a cognitive level, which may influence values and knowledge. The influence that acculturation has on the values may influence attitudes and beliefs about social relations, cultural customs and traditions, gender roles, and attitudes and ideas about health. Knowledge may be influenced by acculturation in the manner in which we recognize or know about culture-specific information, such as names of historical figures belonging to the culture of origin and the dominant culture and the historical significance of culture-specific activities. Lastly, cultural identity has been proposed as a dimension of acculturation. Cultural identity refers to the attitudes an individual has about his or her culture, such as feelings of comfort, pride, or shame toward the culture of origin or the host culture.

Theoretical and Conceptual Frameworks

Medicine and psychiatry had a major influence in the study of acculturation, and early theories of acculturation assumed a clinical viewpoint on matters corresponding to culture contact and change. Much of the early work on acculturation focused on anxiety occurring during cross-cultural transition.

Berry's Model of Acculturation

Most of the current literature uses Berry's model of acculturation to distinguish between the four models (assimilation, separation, integration, and marginalization) of acculturation.

Assimilation

The assimilation model of acculturation characterizes individuals that are highly acculturated; assimilated individuals strongly identify with the dominant or host culture, resulting in the loss of the original cultural identity. The assimilation model of acculturation has come to be known as *cultural shift.* Assimilated individuals that no longer identify with their culture of origin may behave in a manner that no longer reflects the behaviors of the original culture. For

example, assimilated individuals may no longer speak the native language, listen to native music, take part in native dances, or follow the native culture's dating process. Along with behavioral changes, assimilated individuals shift their beliefs, values, and attitudes to match those of the dominant or host culture.

Separation

In the separation model of acculturation, also referred to as *cultural resistance*, an individual will maintain a strong identification with the culture of origin and does not accept the behaviors, attitudes, beliefs, or values of the dominant or host culture. Although an individual may be presented with opportunities to acculturate, the individual consciously chooses to maintain an allegiance with the culture of origin. In this model the individual only displays the behaviors, attitudes, beliefs, and values of the culture of origin.

Integration

The integration model of acculturation, also referred to as *cultural incorporation* and *biculturalism*, is exactly what the term implies. The integration model is a merge and combination of two cultures: the culture of origin and the new dominant or host culture. Individuals in this model may successfully display behaviors, attitudes, beliefs, and values from both cultures. Individuals in this model identify with both cultures and have a level of comfort within both cultures.

Marginalization

The fourth model of acculturation is marginalization; the marginalization model is described as a rejection or nonacceptance of the behaviors, attitudes, beliefs, and values of both the culture of origin and the new dominant or host culture. It is important to keep in mind that a marginalized individual can maintain cultural competence with both groups and have marginal traits as well. Additionally, a degree of acculturation or identification with both cultures must occur before marginalization takes place.

Acculturative Stress

One potential outcome or response of acculturation is acculturative stress, which may result from differences in language, perceived cultural incompatibilities, and cultural self-consciousness. Some stress behaviors that have been associated with acculturation are anxiety, depression, feelings of alienation, and identity confusion.

It is important to note that some research has shown that acculturative stress is not related to the level of acculturation. Thus, one cannot assume that less acculturated individuals experience more acculturative stress than more acculturated individuals. Scholars have also suggested that acculturative stress can stem from the demands to maintain or learn one's cultural heritage while at the same time feeling pressured by the dominant culture to assimilate.

Miguel Ángel Cano

See also Bilingualism; Hispanic Americans; Identity Development; Self-Determination; Social Development

Further Readings

Berry, J. W. (1980). Acculturation as varieties of adaptation. In A. M. Padilla (Ed.), *Acculturation: Theory, model, and some new findings* (pp. 9–25). Albany: State University of New York Press.

Branton, R. (2007). Latino attitudes toward various areas of public policy: The importance of acculturation. *Political Research Quarterly*, *60*(2), 293–303.

Chung, P. J., Travis, R., Jr., Kilpatrick, S. D., Elliott, M. N., Lui, C., Khandwala, S. B., et al. (2007). Acculturation and parent-adolescent communication about sex in Filipino-American families: A community-based participatory research study. *Journal of Adolescent Health*, *40*(6), 543–550.

Kim, B. S., & Abru, J. M. (2001). Acculturation measurement: Theory, current instruments and future directions. In J. G. Ponterotto, J. M. Casas, L. A. Suzuki, & C. M. Alexander (Eds.), *Handbook of multicultural counseling* (pp. 394–424). Thousand Oaks, CA: Sage.

Unger, J. B., Ritt-Olson, A., Wagner, K., Soto, D., & Baezconde-Garbanati, L. (2007). A comparison of acculturation measures among Hispanic/Latino adolescents. *Journal of Youth and Adolescence*, *36*(4), 555–565.

Ward, C. (1996). *Acculturation.* Thousand Oaks, CA: Sage.

Adult Learning

Adult learning is a complex phenomenon. Although there are many commonalities between how adults

and children learn, development and change that take place across the life span play a definite role in understanding why adults learn, as well as how they learn. In educational psychology, an understanding of adult learning and the adult learner is crucial to a commitment to lifelong learning. Major challenges facing those who work with and study adult learners include questions about the extent to which knowledge is discovered or constructed and where one locates oneself in terms of focusing on individual and social aspects of learning.

History

In the early 1900s, the psychology of adulthood and aging received scant attention, largely due to the strong influence of psychologists such as John Watson, Sigmund Freud, and Jean Piaget, whose work emphasized the view that adulthood could be understood as simply an extension of development or learning in the early years of life. However, as psychologists such as Erik Erikson, G. Stanley Hall, Charlotte Bühler, and Sidney Pressey began to extend their theories and research to address the adult years, a greater understanding of the psychology of adulthood began to emerge.

Probably the first major study of adult learning was published by E. L. Thorndike and his colleagues in the 1928 book *Adult Learning*. As a seminal effort to provide empirical evidence related to learning in adulthood, these authors concluded that learning ability peaks at about age 45, rather than age 20 as previously believed. This study set in motion an effort to understand adult learning ability; this effort has continued to grow over nearly eight decades. Today, most of the research and writing on adult learning come from three fields: psychology, adult education, and gerontology. The field of psychology has offered a necessary, though not sufficient, foundation for understanding adult learning. Adult education and gerontology, through their professional literature, have also made important contributions to learning in adulthood. Although there is some overlap across these areas, each field brings a different framework and importance to the understanding of adult learning.

Participation in Adult Learning

One of the most extensively studied areas of adult learning relates to the nature of participation in adult learning. This involves three questions: Who participates in adult learning? Why do adults engage in learning? What are some of the factors that deter or limit adults from participating in learning? In 1965, William Johnstone and Ramon Rivera reported on a major national study of adult learning participation and found that 22% of all adults in the United States participated in some form of learning activity during the previous year.

Beginning in 1969, the U.S. National Center for Education Statistics began to collect data about participation in adult education. These studies have been conducted every several years and have offered insight into some major trends. Unfortunately, because different data collection procedures and definitions were used at different times, direct comparison across studies is not feasible. Nonetheless, it is possible to identify certain trends in participation.

Data from 2000–2001 indicate that 46% of all adults participated in some form of adult education activity during the previous 12 months. This compares with participation rates of 40.2% in 1995 and 33% in 1991. In terms of demographic breakdown, those in the 41-to-50 age group had the highest participation rate (55%), followed by the 16-to-30 and 31-to-40 age groups (53% each). Female participants outnumbered males (49%–43%), and in terms of race/ethnicity, Whites had the highest participation rate (47%), followed rather closely by African Americans (43%) and Hispanics (42%). As might be expected, those with higher educational attainment and income levels also had the highest participation rates. Finally, people who were employed during the previous 12 months had over twice as high a participation rate as those unemployed during the same period (54% compared with 25%).

The way in which participation is defined has an impact on the actual rates of participation. In the data cited in the previous paragraph, participation was defined as some form of adult education class. However, in the early 1970s, Allen Tough used a structured interview process to assess involvement of adults across the entire range of learning activities. In doing so, Tough found that as many as 90% of the participants in his study engaged in some form of learning activity over the previous year. Even more important is that he found the vast majority of learning projects (68%) were planned by the learners themselves, as opposed to a teacher/instructor, tutor, or nonhuman resource such as technology. Tough used the metaphor of an iceberg to illustrate what he found: Most

adult learning lies beneath the surface and is not easily visible to those who only study participation in organized courses or other activities. This finding played a key role in stimulating research on self-directed learning, which became one of the most widely studied topics in the adult education literature of the last decades of the 20th century.

To a large degree, participation in adult learning is linked to life transitions. These transitions often serve as "triggers" for adults to recognize that learning can help them negotiate such transitions. In today's world, job-related transitions (e.g., job loss, promotion, new responsibilities, retirement) are the most frequently identified reasons for participation. Other examples of transitions that can trigger the need for learning include family issues like marriage, divorce, and parenthood; health issues such as wellness, coping with a life-threatening illness, or being diagnosed with a chronic condition; enrichment opportunities such as leisure, art, or religion/spirituality.

One particularly influential study is Cyril Houle's *The Inquiring Mind*, which was originally published in 1961. Houle interviewed 22 adults deemed to be active learners and from these interviews, he identified a typology of orientations toward learning. Goal-oriented learners viewed learning as instrumental to achieving some other purpose; in other words, learning was seen as a means to another end. Activity-oriented learners sought out learning activities for their social value, as a way to meet new people and socialize. Learning-oriented adults were identified as those who engaged in learning for its own sake.

Many factors can serve as barriers or deterrents to participation in adult learning. These have been conceptualized according to several different categories. In essence, the major factors that limit participation are (a) reasons linked to the life circumstances of adults, which are often outside the control of the person, such as lack of time, money, transportation, and family responsibilities; (b) reasons related to institutional policies and practices that limit participation, such as scheduling of classes, information about offerings, limited offerings, and policies that discriminate directly or indirectly against adult learners; and (c) reasons related to attitudes and values of learners, such as low self-concept, fear of failure, negative past experiences, and lack of interest. In the 1980s and early 1990s, Gordon Darkenwald and his colleagues, Craig Scanlon, Thomas Valentine, and Elisabeth Hayes, reported on the development of several forms of the Deterrents to Participation Scale. This instrument has made it possible to isolate and identify many of the factors that contribute to nonparticipation in adult learning.

Intelligence, Memory, and Cognition

Central to the study of adult learning, especially in the realm of psychology and gerontology, has been the body of scholarship on the related areas of intelligence, memory, and cognition. Paul Baltes has described the aging mind as having both potentials and limits, resulting in age-related gains and losses. Whereas some areas of functioning show decline, other areas remain stable and, in some cases, show improvement with age.

To understand changes in intelligence, memory, and cognition over the adult life span, it is important to recognize that much of what has been reported is influenced by the types of research design that have been utilized. Cross-sectional studies measure different age cohorts at a single point in time. These studies make it possible to look at age *differences* on the variable(s) being studied; however, they do not accurately describe age *changes*. Longitudinal studies, on the other hand, measure the same cohort over time, making it possible to study changes that take place over time, but not cohort differences. Most of the early studies on intellectual functioning relied on a cross-sectional approach, and in such cases, researchers often incorrectly identified age-related declines when, actually, what they were observing were cohort differences. As is easy to picture, longitudinal research is difficult to carry out because it requires researchers who can envision and remain committed to a study over many years and even decades. Other problems with longitudinal studies are attrition of participants and instrument decay resulting from changes in the social context that can make earlier instruments irrelevant over time.

One way to minimize the limits of cross-sectional and longitudinal designs is to use an approach that combines both approaches in a single study. Here, longitudinal data are collected over time with a single cohort. During each measurement, however, a new cohort of younger participants is added. Eventually, this design will generate enough data to address age changes over time as well as cohort differences. Perhaps the most influential study of this type is the Seattle Longitudinal Study, developed by K. Warner

Schaie. Begun in 1956, with additional cohorts added every 7 years, the Seattle Longitudinal Study has focused on five mental abilities: (1) *verbal meaning*, which refers to the ability to understand ideas as expressed in words; (2) *spatial orientation*, the ability to visualize, manipulate, and perceive connections among objects; (3) *inductive reasoning*, the ability to recognize or make sense of new concepts and the ability to analyze and solve problems and situations; (4) *numeric ability*, which refers to understanding numbers and figures and the speed and accuracy with which a person can solve numerical problems; and (5) *word fluency*, involving ability to recall words in writing and speech. Basically, this study has presented evidence that in normal aging, there is little or no discernable decline in primary mental abilities until the mid-to-late 60s and this decline is slow until the 80s.

An area of debate related to intelligence in adulthood centers on whether intelligence is a general factor, as is typically defined in IQ tests, or whether there are different kinds of intelligence that account for a wide range of abilities. The view that there is more than one type of intelligence was introduced by Raymond Cattell and John Horn, who distinguished between *fluid* and *crystallized* intelligence. Fluid intelligence is a biologically based form of intelligence that is innate and involves reasoning ability. Crystallized intelligence, on the other hand, is largely dependent on education and experience. Thus, in this view, there is evidence that whereas fluid intelligence is characterized by age-related decline, crystallized intelligence, by building on past experience, typically increases over the life span.

In recent years, two theories have proposed that intelligence comprises multiple factors. Robert Sternberg has proposed a triarchic theory of successful intelligence, which holds that intelligence comprises a mix of analytical, creative, and practical abilities. The first of these is the more traditional view of academic intelligence. Creative intelligence centers on how well one addresses new and unfamiliar situations. Practical intelligence has to do with how effectively one is able to adapt to and solve everyday problems. A second approach to the multifactor view of intelligence is Howard Gardner's theory of multiple intelligences. According to Gardner, there are eight intelligences that address a wide range of abilities. These intelligences include linguistic, logical-mathematical, spatial, musical, bodily-kinesthetic, and two forms of personal intelligences that involve understanding oneself and others.

The eighth intelligence, naturalistic, has been added recently and is based on an understanding of the natural environment. The theories proposed by Sternberg and Gardner have relevance to adult learning because they recognize and value types of intelligence that extend beyond the more traditional IQ-based approach to intelligence. Because adult learning typically has a practical bent, emphasizing different kinds of abilities, the notion of multiple intelligence holds much potential for future research and practice with adult learners.

Memory is closely linked to intelligence. It involves the acquisition, retention, and recall of information. Although there are many schema for distinguishing among different kinds of memory, the distinction between short-term and long-term memory serves well to illustrate how information is retained and recalled in adulthood. *Short-term memory*, which typically covers a period of 10 to 30 seconds, can be further broken down into primary memory and working memory. Primary memory is more passive and involves holding information for immediate recall (e.g., remembering a phone number long enough to go to the phone and dial it or remembering information on a road sign when driving long enough to follow the desired direction). Working memory is more active and centers on the amount of information that can be held in memory long enough to perform some other operation on it. Research evidence suggests that whereas the changes in primary memory are small and gradual over time, there is a major decline in working memory with age. This decline has been attributed to a host of factors, including (a) a decline in mental energy that can result in overloading with increasingly complex tasks, (b) a weakened ability to use strategies related to working memory, and (c) a decline in speed of memory processing.

Long-term memory involves how facts are stored (semantic memory) and the ability to recall events from the past (episodic memory). Semantic memory is typically stable into the 70s and then declines gradually. On the other hand, episodic memory tends to decline with age although it is possible to compensate for some of this loss. Declines in long-term memory have variously been attributed to how material is acquired, how it is retrieved, and how fast it can be processed. At the same time, some researchers have suggested that memory training activities can help adults retain the ability to use knowledge, strategies, and skills.

Cognition involves all forms of knowing and awareness, including, but not limited to, information

processing, problem solving, perceiving, abstract reasoning, and judging. Much of the work on cognition in adult learning acknowledges the work of Jean Piaget as a starting point; however, because Piaget focused primarily on early development, subsequent work on cognitive development in adulthood has attempted to move beyond Piaget's original ideas. One model of cognitive development that is often adapted to the adult learning context is William Perry's intellectual development scheme. Based on data from male Ivy League college students, Perry found that as learners develop, they move from *dualistic* thinking, where "right and wrong" answers are presented by authorities, to *relativistic* thinking, where understanding the context is as important as the knowledge itself. An important response to Perry's scheme is the work of Mary Belenky, Blythe Clinchy, Nancy Goldberger, and Jill Tarule, who looked at "women's ways of knowing." They identified five categories of knowing, ranging from "silence," where women lack voice and are subject to what is expected from authority figures, to "constructed knowledge," where women perceive themselves able to create knowledge and to recognize all knowledge as contextual. This important study offered evidence of ways in which the experiences of women can differ from those of men.

A question that lies at the heart of cognition centers on whether knowledge is discovered or constructed. Whereas many views of cognition emphasize knowing as an internal process of uncovering knowledge that goes on within the individual learner, the literature on adult learning has increasingly focused on the social context in which learning takes place. This approach to understanding knowledge is often referred to as *situated cognition.* In situated cognition, knowledge cannot be separated from the context in which learning takes place. Thus, learning involves the construction of knowledge within the social milieu in which it occurs. Because it emphasizes learning in social context, situated cognition is characterized as having an inherently social or political element to knowledge and a connection to the importance of power in relation to cognition.

Theories and Concepts in Adult Learning

There is no single theory or model that explains adult learning. This should come as no surprise, as adults engage in learning for myriad reasons and in virtually unlimited settings. Adults become involved in learning for such wide-ranging purposes as work/career-related learning and training, literacy and basic skills development, earning credentials or degrees, social change or civic education, and personal growth, enrichment, and enjoyment. Adult learning takes place in such diverse settings as colleges and universities, public schools, business and workplace settings, health and human services agencies, churches, civic organizations, and government agencies, including the military. Although there is no single theory or model that can fully explain adult learning, there are nonetheless several theories or concepts that illustrate the scope of what is known about adult learning. Five of these are described in the following sections.

Andragogy

Andragogy is a concept that can be traced back to the 19th century. However, it came into prominence in adult learning during the late 1960s and early 1970s through the work of Malcolm Knowles. Knowles was looking for a framework to distinguish adult learning from learning in childhood (pedagogy). Later, Knowles suggested that pedagogy and andragogy are not limited to children and adults, respectively, but are more matched to the maturity and experience of the learner in a given setting. Though andragogy is sometimes described as a theory, it is probably more accurate to describe it as a set of assumptions about how people learn. According to andragogy, as learners mature,

1. their self-concept moves from being dependent to increasing levels of self-directedness;

2. the role of the learners' experience becomes an increasingly valuable resource, and adult learning is optimized when learners are able to tap into their experience;

3. readiness to learn in adulthood is increasingly based on real-life needs and situations;

4. there is a shift from learning for future application toward learning to address immediate needs;

5. intrinsic motivators become increasingly more important than extrinsic ones; and

6. before learning something, adults typically need to know why they need to learn it.

Although some critics have challenged andragogy because its focus is largely on the individual learner and does not directly address the social context in which learning takes place, andragogy retains an important place as a set of practices that have value when working with adult learners.

Self-Directed Learning

Closely related to the first assumption of andragogy, self-directed learning emerged in the early 1970s as one of the most systematic areas of research and scholarship on adult learning. Much of this interest grew out of the research of Allen Tough, which was mentioned earlier in this entry, on adults' learning projects. Subsequently, two individuals who have played a key role in developing this work, both through their own writings and through their many doctoral graduates, are Roger Hiemstra and Huey Long. Though there are many definitions, models, and conceptualizations of self-directed learning, it is essentially where the learners assume primary responsibility for and control over their learning. In 1991, Ralph Brockett and Roger Hiemstra synthesized ideas from several previous authors to present a model that describes self-direction as the product of two factors: the teaching-learning situation and internal characteristics of the learner. Out of this body of research, several ideas have emerged: first, self-directed learning is the most frequent way in which adults choose to learn; second, self-directedness has a strong connection to how learners feel about themselves as learners; third, several personality and social characteristics seem to have a connection to self-directedness; and finally, research on self-directed learning has probably contributed to a more holistic understanding of adult learners' potential.

Transformative Learning

Experience lies at the heart of adult learning. In most approaches to learning in adulthood, experience is acknowledged to be an important resource for the learner. However, it is not merely having an experience that matters; rather, it is the way in which the learner makes meaning of the experience and is changed by the experience that is important. In the 1970s, Jack Mezirow reported on a study of women who had returned to college. He described a process similar to consciousness-raising that took place for many of the women. Instead of simply acquiring knowledge, the experience of returning to college often resulted in a transformation that transcended the college experience and led to a redefining of the self. Initially, Mezirow referred to this process as *perspective transformation* and laid the foundation for a theory that has evolved over the ensuing three decades. Mezirow cited several key influences including Jürgen Habermas (critical theory), Roger Gould (psychoanalytic psychology), Paulo Freire (conscientization), and Thomas Kuhn (paradigm shifts). Through exchanges with various scholars who challenged aspects of Mezirow's theory, particularly those related to the emphasis on social change, Mezirow and these scholars refined the theory over time and began to use terms such as *transformation theory, transformational learning*, and *transformative learning* to describe this concept.

Transformative learning typically begins with a disorienting dilemma that makes it necessary for a person to examine existing assumptions and frames of reference. Examples of a disorienting dilemma include job loss, diagnosis of an illness such as cancer or diabetes, loss of a spouse or partner through death or divorce, or some sort of spiritual awakening. The dilemma, by its nature, redefines this aspect of one's life in a way that cuts across one's roles, responsibilities, and identity. Transformative learning involves a process of challenging assumptions in a way that helps one redefine oneself. An example of this would be an adult who returns to higher education and finds oneself immersed in study in a way that the person becomes socialized in an academic or professional field, and this new identity emerges as a vital part of who the person becomes.

In recent years, the literature on transformative learning has snowballed as various scholars have expanded on Mezirow's original ideas. John Dirkx, for example, has distinguished among four types of transformational learning. He describes Mezirow's view as a "cognitive-rational" perspective, which shares constructivist theoretical underpinnings with Freire's "emancipatory" approach but differs from Freire by emphasizing the process of reflection and rational thought. A third approach is the "developmental" perspective of Larry Daloz, which is holistic and contextually based and emphasizes how individuals negotiate developmental transitions in their lives. Finally, Dirkx describes a "spiritual-integrative" approach that extends beyond the rational approach to focus on feelings and images emerging from soul-based learning. The point here is that Mezirow's work is seminal to

understanding transformative learning, and subsequent work by various individuals, including work published in such sources as *Adult Education Quarterly* and the more recent *Journal of Transformative Learning*, demonstrates that this topic has assumed a central place in current scholarship on adult learning.

Humanism and Behaviorism

As with the field of educational psychology in general, humanism and behaviorism have held an important place in the conceptual foundations of adult learning. Humanism, with its view of human nature as basically good and the belief that individuals have virtually unlimited potential for growth, is particularly attractive to those seeking to bring out the best that adult learners can achieve. In humanistic adult education, Carl Rogers and Abraham Maslow are among the seminal thinkers. The humanist perspective underlies many of the tenets of andragogy and self-directed learning and is most often reflected in adult learning activities directed toward personal growth or improvement.

Behaviorism has also had a strong influence in the area of adult learning. With emphasis on reinforcement, learning for mastery, and helping people achieve competency in what is being learned, behaviorism has been especially influential in settings where performance and measurable outcomes are of primary importance. The influence of behaviorism is perhaps strongest in adult learning situations such as training in business and industry and adult literacy settings, where achieving measurable outcomes is deemed central to the success of learning. It can also be found where learning is designed to help people change specific behaviors, such as weight loss, smoking cessation, and substance abuse learning efforts.

Critical and Postmodern Perspectives

A major challenge to historically influential, learner-centered adult learning theories has, in recent years, come from critical and postmodern perspectives. Here, the focus is less on individual growth and development and more on issues of power and voice and the sociocultural context of learning. Sharan Merriam, Rosemary Caffarella, and Lisa Baumgartner have noted that major themes of these theories include race, class, and gender; power and oppression; and knowledge and truth. Because critical and postmodern theories emphasize the context in which learning takes place and the idea that

knowledge is socially constructed outside of the individual learner, these approaches are especially compatible with learning for social change. However, those who approach learning from critical and postmodern orientations tend to reject or minimize the importance of theories and research derived from a psychological orientation. For those interested in the psychology of adult learning, critical and postmodern theories pose a challenge in terms of the degree to which emphasis should be placed on individual and social dimensions of learning.

Ralph G. Brockett

See also Cognitive View of Learning; Crystallized Intelligence; Fluid Intelligence; Intelligence and Intellectual Development; Learning; Lifelong Learning; Long-Term Memory; Multiple Intelligences; Older Learners; Short-Term Memory; Triarchic Theory of Intelligence; Working Memory

Further Readings

Birren, J. E., & Schaie, K. W. (Eds.). (2001). *Handbook of the psychology of aging.* San Diego, CA: Academic Press.

Bjorklund, B. R., & Bee, H. L. (2007). *The journey of adulthood* (6th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

Hoare, C. (Ed.). (2006). *Handbook of adult development and learning.* Oxford, UK: Oxford University Press.

Jarvis, P. (2006). *Towards a comprehensive theory of human learning.* London: Routledge.

Merriam, S. B., Caffarella, R. S., & Baumgartner, L. M. (2007). *Learning in adulthood* (3rd ed.). San Francisco: Jossey-Bass.

Schaie, K. W. (2005). *Developmental influences on adult intelligence.* Oxford, UK: Oxford University Press.

Tennant, M. (2006). *Psychology and adult learning* (3rd ed.). London: Routledge.

AFFECTIVE DEVELOPMENT

See ATTACHMENT

AFRICAN AMERICANS

This entry provides a summary of issues related to African Americans. It begins with a brief overview of

the history and geographical distribution of African Americans in the United States. This is followed by brief descriptions of attitudes affecting African Americans, including their attitudes toward the majority culture, religion and politics, racial identity, and African Americans in society. Next is a summary of the schooling experiences of, and educational outcomes for, African Americans and an overview of some of the major explanations for the achievement patterns of African American students. Despite decades of educational research, African Americans continue to have achievement levels that are substantially below their White and East Asian counterparts, on average, and this achievement gap has been the subject of considerable research in the field of educational psychology.

African Americans in the United States

At 12.3% of the population, African Americans are the second largest minority group in the United States, behind Hispanics who make up 12.5% of the U.S. population. Also commonly referred to as Black Americans, African Americans are still the single largest *racial* minority group in the country. Many Black biracial individuals and foreign-born Blacks also self-identify as African American. However, both of these groups are quite small, constituting less than 1% and 5% of the Black population, respectively. Black immigrants to the United States are from three primary regions: the Caribbean, Latin America, and sub-Saharan Africa.

In the early 1600s, some of the first Africans were brought to the United States as indentured servants. However, by the mid-1600s, the status of Africans in the United States had been changed to slaves. By the end of the legal slave trade in 1808, more than 500,000 Africans had been brought into the United States, primarily to work on the plantations in the southern states. Until the Great African American Migration (1910–1920), 90% of the African American population lived in the South, and less than 25% lived in urban areas. This early 20th-century migration of African Americans eventually yielded settlement patterns that are reflected in the United States today, with substantial concentrations of African Americans in the urban centers of the northeastern and midwestern United States.

The South still has the largest concentration of African Americans in the country, with more than 50% of the Black population. Nonetheless, African Americans constitute less than 20% of the South's population. African Americans make up about 11% of the populations in the Northeast and Midwest and about 5% of the population in the West. The top 10 residential areas for African Americans include the following metropolitan areas, in descending order: New York City, Chicago, Washington, D.C., Atlanta, Detroit, Philadelphia, Los Angeles/Long Beach, Houston, Baltimore, and Dallas. Although there have been declines in segregation over the past three decades, African Americans still live in highly segregated neighborhoods, and the metropolitan areas with the largest concentrations of African Americans also have the highest levels of residential segregation. Affluent African Americans also live in areas with high levels of residential segregation, and children experience the highest segregation levels at home and in school.

Attitudes Affecting African Americans

African Americans report a substantial amount of mistrust of Whites, and a majority of African Americans report experiencing racial discrimination at least once a year. African Americans report marked mistrust of the police and are more likely than other ethnic groups to believe in conspiracy theories. Thus, many African Americans believe that the majority culture does not want positive outcomes for their group. The increasing numbers of legal challenges to affirmative action, alongside media accounts of police brutality, reports of discrimination in housing and bank loans, and concerns about the disenfranchisement of Black voters in recent national elections, are seen as providing support for this hypothesis.

There is some evidence that African Americans' beliefs are not wholly without merit. Although overt negative attitudes toward African Americans have decreased in the broader society, recent research suggests that racist attitudes have become more covert over time and that aversive (unconscious) racism still plays a major role in the lives of African Americans. Thus, although fewer than 5% of Americans endorse negative stereotypes of African Americans, social psychology research reveals differential responses to African Americans in several circumstances. For example, in bystander intervention studies, individuals respond equally to Whites and Blacks when an emergency occurs *only if* they are the only witness; they are less likely to assist Blacks when there are multiple

witnesses. Also, people report the same number of negative characteristics for Whites and Blacks but report significantly more positive characteristics for Whites. Similarly, in evaluating job candidates, Black and White candidates with comparable weak and moderate skills are rated equally, but White candidates with strong skills are rated substantially higher than Black candidates with equally strong skills.

African American racial identity attitudes have also received considerable research scrutiny. Black racial identity was first conceptualized as a stage theory, with African Americans moving from an anti-Black stage to a pro-Black stage, with concomitant increases in self-concept. Recent theorizing conceives of Black racial identity as consisting of a set of multidimensional attitudes. Studies have provided support for several Black racial identity attitudes, including low salience attitudes, high salience attitudes, pro-Black attitudes, anti-Black attitudes, anti-White attitudes, and multicultural attitudes. Recently, researchers have demonstrated generalizable racial identity profiles in the population, although it is not yet clear how these will affect functioning. Several scholars have suggested that in the social context of the United States, many African Americans will have some minimal level of negative attitudes toward Whites.

Religion continues to play a major role in African American life, and the church has been one of the most important forces in the African American community. Estimates indicate that there are more than 60,000 Black churches in the United States, and African Americans have considerable membership numbers in several Christian and Islamic denominations. From a religious point of view, African Americans are generally conservative. However, on the political spectrum, African Americans are generally aligned with the Democratic Party. Initially strong supporters of Abraham Lincoln and the Republican Party after the American Civil War, African American political allegiance shifted as the Republicans compromised with southern states on issues of civil rights. For much of the past 30 years, more than 80% of African American voters have been affiliated with the Democratic Party, and the 109th Congress included 43 African Americans (1 senator and 42 representatives), all of whom were Democrats.

Educational Issues

African Americans comprise approximately 17% of the enrollment in public schools. However, they constitute about 20% of the students in special education, 30% of the students in vocational education, 23% of the students in alternative schools, and only 12% of the students in gifted and talented programs. African Americans also make up 10% of the private school enrollment. Overrepresentation in special education is greatest in the categories of mental retardation, developmental delay, emotional disturbance, deaf-blind, autism, and multiple disabilities. In contrast with the school population nationally, which is concentrated in suburban schools, more than 50% of African American students attend urban schools. On average, African Americans attend schools of lower quality with higher levels of segregation than other groups, even though it is 50 years after the *Brown v. Board of Education of Topeka, Kansas*, decision. In this case, the U.S. Supreme Court indicated that schools that were separate were inherently unequal. African American students report concerns about violence and the availability of drugs, alcohol, and weapons in the schools that they attend in substantially greater percentages than do other ethnic and racial groups.

One of the more persistent problems in the education arena is the achievement gap between African Americans, Latinas/os, and Native Americans on the one hand, and Whites and Asian Americans on the other. On average, African American students enter elementary school with weaker math, vocabulary, and reading skills than their White counterparts, even after controlling for parents' education levels, and this gap in achievement widens from Grades 1 to 12. Significantly fewer African American preschoolers and kindergartners can identify all the colors and alphabet letters, and twice as many African Americans in this age group are diagnosed with learning disabilities compared with their White counterparts. The event dropout rate (i.e., the percentage of students who dropped out of high school in a given year) for African American students is about 6%, and the status dropout rate (i.e., the percentage of individuals in the population from a certain age group who are not enrolled in school and have not earned a high school diploma) for African Americans aged 16 to 24 is approximately 13%. Dropout rates for African Americans declined substantially from the 1970s, but they stabilized in the late 1990s. In some urban districts, African American graduation rates are below 50%.

Data from longitudinal studies of school-age adolescents indicate that African American students miss

more days of school than the aggregate U.S. student population and have the highest suspension and expulsion rates. African American students also report spending more time watching television on weekdays and weekends. African American males are overrepresented in both incarcerated youth and youth on probation. These disparities are also reflected in both educational and occupational attainment. African Americans have lower college enrollment and graduation rates than do White and Asian students, and the percentages of African American workers decrease as one moves from clerical up to professional positions.

Explanations of African American Achievement Patterns

Several explanations have been advanced for African American underachievement, although it is likely that no single factor provides a complete explanation of this complex issue. Many of the initial arguments focused on environmental deficits. For example, it was assumed that African American homes were culturally deficient in ways that precluded academic achievement. However, differences related to academic achievement (e.g., quantity and quality of language in the home) were related more closely to class than to racial group. Deficits in the segregated schools that African Americans attended were also identified as a major concern, and one of the major legal accomplishments of the civil rights era was the *Brown v. Board of Education of Topeka, Kansas*, decision in 1954. This decision led to decades of desegregation plans by school districts, some of which are still in place.

Another long-standing argument has been based in biology and genetics—that is, African Americans have lower scores on measures of *g* (general intelligence) and consequently lower academic achievement. However, this explanation has been criticized for ignoring data on the increase in IQ scores over the past century (the Flynn effect), the reciprocal relationship between effective schooling and IQ, and the differential contributions of IQ to the variance in achievement across socioeconomic groups. There have also been suggestions that African Americans have different cultural styles than Whites and that there is a mismatch between Black cultural styles and the common methods of teaching in the school system. However, this suggestion has not been supported in empirical studies, in part due to the failure of researchers to adequately define and operationalize culture.

Socioeconomic status has also been put forward as a reason for African Americans' underachievement, as there is a moderate relationship between poverty and academic achievement. African Americans are one of the poorer groups in the United States, with 25% of African American adults and 33% of African American children living below the poverty line. In addition, more than 40% of the African American population live in households with annual incomes of less than $25,000, and African Americans comprise about 40% of the homeless population. Black males also have the highest unemployment rate, and only 12% of all African American households report incomes of more than $75,000. Although socioeconomic status plays a role, there are data indicating that the achievement gap is present at all socioeconomic levels.

Although there are consistent positive relationships between academic achievement and several psychosocial variables across racial and ethnic groups, including academic self-concept, academic self-efficacy, self-regulation, motivation, and future time perspective, there have been few studies of these constructs in African American populations, and the studies that exist have been of limited utility in explaining the achievement gap. African American students consistently report higher self-concept scores and educational and occupational expectations, despite lower academic achievement. There is a growing consensus that the motivation and future orientation of African Americans may be affected by the marginalized status of African Americans in society. African American role models in the public sphere are more frequently entertainers and athletes than academics.

Another set of explanations for African American achievement is psychosocial and implicates African Americans' collective or social identity, also referred to as their *reference group orientation*. Perhaps the best-known of these explanations is the cultural ecological theory proposed by the late educational anthropologist, John Ogbu. This model has been used to explain achievement differences across racial and ethnic groups around the world, including Australia, Great Britain, Japan, and the United Kingdom, and has also been used to explain achievement differences among racial and ethnic groups in the United States. In brief, cultural ecological theory suggests that one of the ways in which African American educational success is compromised is due to the group eschewing attitudes and behaviors that are conducive to educational attainment if the attitudes and behaviors are seen as "acting White."

Another explanation involving reference group orientation is the stereotype threat phenomenon proposed by Claude Steele, a social psychologist. This argument focuses on the pervasiveness of stereotypes in the populations and suggests that the strong negative stereotype about African Americans' intellectual capacities in the population can have a detrimental impact on African American performance in situations (e.g., academic evaluations of performance) where the stereotype is salient. Some researchers suggest that African Americans who care the most about doing well and have a strong sense of bonding to their racial group are potentially at greater risk for stereotype threat. Although much of the early research on stereotype threat was conducted with college-age samples, recent studies have demonstrated the negative impact of stereotype threat on school-age African Americans.

More recently, researchers have found that African American college students who are high in race-based rejection sensitivity are less likely to seek assistance in predominantly White institutions and more likely to be socially isolated, potentially decreasing their chances of persisting until graduation. Finally, researchers who study racial and ethnic identity in African Americans have hypothesized that some racial identity profiles may be more compatible with schooling outcomes than others, although there are limited data in support of this hypothesis. In addition, there are data that indicate that African Americans do not have to abandon their racial identity to be successful in school.

Several other theoretical perspectives implicate attitudinal and personal identity variables in academic performance. Self-worth theory suggests that students who are concerned with protecting their academic self-concept may *choose* not to study for examinations or engage actively in learning, thus providing a clear rationale for poor performance (i.e., not studying), rather than studying and taking the chance that they fail and be perceived as unintelligent. Another motivational perspective suggests that African American adolescents may have to choose between belonging to their racial group or being high achievers, leading many capable African American students to resist showing their true academic potential (i.e., *I can, but do I want to?*). Other researchers have reinvigorated the literature on teacher expectation effects by demonstrating that students can recognize differential behaviors by teachers toward high and low performers from the early elementary years onward. These findings suggest that the differential treatment of students perceived as more capable and less capable has direct effects on students' classroom functioning. The findings also indicate that minority students, who can recognize that they are members of stigmatized groups from the elementary school years, may be especially sensitive to the messages that teachers are communicating.

Health Disparities

As with the achievement gap, there are disparities in health between African Americans and White Americans across the life span. African Americans are either first or second with regard to infant mortality rates, low birthweight infants, and teenage pregnancies. African Americans also have the highest death rates in the United States, both generally and from specific causes such as heart disease, stroke, lung cancer, female breast cancer, and homicide. African Americans make up a substantial proportion of males (34.7%), females (60.5%), and children (61%) dying from AIDS. With regard to mental health issues, African Americans have less insurance and are also less willing to seek treatment, resulting in lower access to mental health care. The dearth of African American providers also contributes to an unwillingness to seek treatment.

Research Agenda

African Americans make up a substantial portion of the U.S. population. They are a group with a unique place in the history and current sociopolitical context of the United States, and their academic achievement relative to members of the majority culture represents one of the most intractable problems facing educational psychology. Educational psychologists need to develop a comprehensive research agenda on this issue. Although efforts to close the achievement gap continue apace, it is perhaps time to rethink the approach to African American achievement. Data from the 2005 National Assessment of Educational Progress indicate that Whites and Asian/Pacific Islanders in the fourth and eighth grades had higher reading achievement scores, 22 to 29 points higher, than their Black, Hispanic, and American Indian counterparts. Moreover, these scores have remained relatively unchanged for much of the past decade.

Beginning in 2002, the No Child Left Behind (NCLB) Act required school districts to disaggregate student performance data on the basis of demographic variables, including ethnicity, gender, race, and

socioeconomic status. One result of this requirement was to highlight the disparities in educational performance among groups at the local level, disparities heretofore seen only in national data. NCLB also proposes eliminating the achievement gap by 2014. However, there are still no explanations of the achievement gap that are accepted as definitive by large segments of the research community. Genetic, biological, environmental, demographic, and psychosocial explanations have both proponents and critics, and there are no studies in which combinations of all of these variables have been systematically examined.

Educational research indicates that effective interventions (e.g., *Sesame Street*) often have positive consequences for those who need them as well as for those who could have done without them. Thus, in the absence of interventions targeted specifically at and solely to African Americans, it may be difficult to *close* the achievement gap. However, the achievement gap between African Americans and Whites may not be the only issue that merits attention. An alternative perspective suggests focusing attention at the achievement levels that African Americans attain. From this perspective, educators should be concerned with ensuring that all African Americans attain at least basic proficiency in the core academic subjects and the skills to pursue advanced educational opportunities.

Research indicates that effective teachers make a difference in student performance and that the impact is cumulative over multiple school years, but it is not clear how these findings relate to African American student achievement. Nor is it clear how teacher education programs need to change to prepare teachers to not only be effective but also convey appropriate messages to African American students. Effective teacher training needs to counter the increased susceptibility of African American students to negative messages conveyed by teachers and help teachers counter aversive racism. More research is needed on the relationships between African American collective identity and psychosocial variables that are proximally related to academic achievement (e.g., self-efficacy, intrinsic motivation, and anticipation of success) and on the mechanisms that are effective in supporting these variables in African American students. Finally, researchers need to assess how much variance in African American student achievement is attributable to racial identity attitudes, oppositional identity, and stereotype threat, and which behaviors and attitudes are compatible with high academic achievement and African American ethnic heritage.

As should be clear, there are many questions that remain unanswered with regard to African Americans and schooling in the United States. Effectively addressing the educational issues related to African Americans will require a focused research agenda and collaboration between researchers and school districts with large numbers of African Americans. Educational researchers concerned with equity, and with African Americans specifically, have a challenging and worthwhile task that requires urgent attention.

Frank C. Worrell

See also Effective Teaching, Characteristics of; Ethnicity and Race; Intelligence and Intellectual Development; National Assessment of Educational Progress

Further Readings

Brody, N. (2003). Jensen's genetic interpretations of racial differences in intelligence: Critical evaluation. In H. Nyborg (Ed.), *The scientific study of general intelligence: Tribute to Arthur R. Jensen* (pp. 397–410). San Francisco: Pergamon.

Ceci, S. J., & Papierno, P. B. (2005). The rhetoric and reality of gap closing: When the "have-nots" gain but the "haves" gain even more. *American Psychologist*, *60*, 149–160.

Graham, S. (1994). Motivation in African Americans. *Review of Educational Research*, *64*, 55–117.

Jencks, C., & Phillips, M. (Eds.). (1998). *The Black-White test score gap.* Washington, DC: Brookings Institution Press.

Lee, J. (2002). Racial and ethnic achievement gap trends: Reversing the progress toward equity? *Educational Researcher*, *31*, 3–12.

Mendoza-Denton, R., Downey, G., Purdie, V. J., Davis, A., & Pietrzak, J. (2002). Sensitivity to status-based rejection: Implications for African American students' college experiences. *Journal of Personality and Social Psychology*, *83*, 896–918.

Neisser, U. (Ed.). (1998). *The rising curve: Long-term gains in IQ and related measures.* Washington, DC: American Psychological Association.

Ogbu, J. U. (2004). Collective identity and the burden of "acting White" in Black history, community, and education. *The Urban Review*, *36*, 1–35.

Polite, V., & Davis, J. E. (Eds.). (1999). *African American males in school and society: Practices and policies for effective education.* New York: Teachers College Press.

Rushton, J. P., & Jensen, A. R. (2005). Thirty years of research on race differences in cognitive ability. *Psychology, Public Policy, and Law*, *11*, 235–294.

Thompson, G. L. (2003). *What African American parents want educators to know.* Westport, CT: Praeger.

Walker, J. E. K. (1999). African Americans. In E. R. Barkan (Ed.), *A nation of peoples: A sourcebook on America's multicultural heritage* (pp. 19–47). Westport, CT: Greenwood Press.

Worrell, F. C. (2005). Cultural variation within American families of African descent. In C. L. Frisby & C. R. Reynolds (Eds.), *Comprehensive handbook of multicultural school psychology* (pp. 137–172). Hoboken, NJ: Wiley.

Yeakey, C. C., & Henderson, R. D. (Eds.). (2003). *Surmounting all odds: Education, opportunity, and society in the new millennium.* Greenwich, CT: Information Age.

AGGRESSION

Aggression is a common problem among schoolchildren and results in negative psychological, educational, and social outcomes for both aggressors and victims. This entry considers this aggression from both sides, that is, the side of the aggressors and the side of their victims. More specifically, it defines the terms *aggression* and *peer victimization* and reviews prevalence estimates of each. It also reviews the consequences of aggression for both aggressors and victims, as well as the antecedents or risk factors for each. This entry then moves beyond these generalities to discuss some of the subtypes of aggression and victimization. Finally, it offers some conclusions that can be drawn from the existing research and describes likely future directions for studying aggression.

Definitions and Prevalence

Aggressive behavior can be defined as any act that is aimed at harming another individual. More specifically, the study of childhood aggression often involves aggressive behaviors among peers, that is, children of similar ages (excluding aggression toward or from adults). Using this definition, attention is placed both on aggressors, who frequently enact aggression toward their peers, and on victims, who are often the targets of aggression by peers. It is important to note that some children may be considered both aggressors and victims; these aggressive-victims often have outcomes and risk factors that are distinct from children who are only aggressors or only victims.

Prevalence estimates of aggressors, victims, and aggressive-victims vary widely across studies because of different measurement strategies (e.g., reliance on children's self-reports or nominations of peers, teacher reports, observations) and criteria for classifying children (e.g., many studies define a child as a victim if they are targeted about once a week or more, but others will consider entire school years or lifetime incidents). Despite this variability across studies, it appears that about 10% to 20% of children can be considered aggressors, 10% to 20% can be considered victims, and 5% to 10% can be considered aggressive-victims. These prevalence estimates are remarkably consistent across countries, so it appears that aggression is a problem among schoolchildren worldwide. It is also worth noting that although these prevalence estimates would suggest that most children (50%–75%) are not directly involved as aggressors and/or victims, most children play some role in aggressive incidents, often serving as assistants or reinforcers to aggressors or as defenders of victims.

Consequences

The substantial prevalence of aggression and victimization is especially alarming when one considers the serious negative consequences of each. Aggressive children are often disliked by their normative (nonaggressive) peers and affiliate with delinquent peers who may solidify and expand the child's antisocial tendencies. Aggressive children are also often disengaged from school, either by their own choice or through negative teacher reactions, suspensions, and expulsions. These negative consequences of childhood are often exacerbated over time, leading to further delinquency, substance use, and school dropout during adolescence and to criminal behavior, poor marital relations, and unemployment/underemployment during adulthood. Of course, these associations are not perfect, and most aggressive children will discontinue, or at least decrease, their use of aggression with time and lead normal, well-adapted lives (in fact, there is evidence that most early adolescents will engage in some antisocial behavior, generally with few long-term consequences). At the same time, these long-term associations suggest that childhood aggression places individuals at increased risk for negative trajectories, and such behavior should certainly not be dismissed as "kids being kids."

As might be expected, victims of peer aggression suffer in numerous ways as a consequence of being abused. Victimization often leads to diminished

self-esteem and increases in internalizing problems (depression, anxiety, social withdrawal). Victims also tend to have poorer academic adjustment, including lower grades, disliking of school, and truancy; these consequences are intuitive if we imagine, as adults, how we would perform at work if we expected that someone might assault us on our next break. Victimization also leads to poor social outcomes, in the forms of having fewer friends, having friendships of poorer quality, and being disliked by most peers. This is unfortunate because the psychological consequences of victimization are diminished for victims who have good social support (e.g., friendships). Although the empirical evidence is limited, that which is available indicates that these negative consequences are long-lasting and persist as increased rates of depression and problematic romantic relationships, for example.

Children who are both aggressors and victims tend to suffer even more serious adjustment difficulties than children who are only aggressors or only victims. The additive risks alone of being both aggressive and victimized suggest negative adjustment, and these aggressive-victims do indeed appear to suffer the short- and long-term consequences of both aggressors and victims. Moreover, there is some evidence that these aggressive-victims suffer even worse outcomes than would be predicted by the additive effects of aggression and victimization. It is unclear if the dual roles of aggressor and victim are especially detrimental, or if the same risk factors that predict children becoming aggressive-victims (e.g., neurological deficits, histories of parental abuse) also contribute to their long-term maladjustment. Nevertheless, these children represent a special cause for concern.

Risk Factors

Given the prevalence and negative consequences of aggression and victimization, researchers have sought to identify factors that place children at risk for enacting and/or receiving aggression.

Predictors of aggressive behavior can be found in both home and peer contexts. Specifically, the home environments of children who enact aggression tend to be characterized by marital conflict and frequent aggression (e.g., domestic violence). Furthermore, aggression is predicted by parenting styles of inappropriate permissiveness or lack of monitoring of children's behavior, negative or rejecting behaviors toward children, and of physical punishment and/or

inconsistent discipline of children's behavior. In the peer context, research has shown that experiences of peer rejection and victimization predict increases in aggression, as do group social norms encouraging aggressive behavior and affiliation with aggressive and/or delinquent peers. It is worth noting that some of these peer-group risk factors for aggression are also consequences of aggression; thus, initial home environment may contribute to children's aggressive behavior, which results in peer relations that further solidify and exacerbate aggressive tendencies.

Victims of peer aggression are more often physically weak, suffer internalizing problems (i.e., depression, anxiety), and have lower self-concept than nonvictimized peers; each of these factors might make children less likely or less able to behave assertively or defend themselves, which may contribute to them being viewed as "easy targets" by potential aggressors. Similar to aggression, risk factors for peer victimization can also be found in both home and peer contexts. Parents who provide little support or responsiveness to their children's needs tend to have children who are more likely to be victimized by peers. Other parenting risk factors differ by gender; for instance, overprotectiveness and enmeshment predict victimization for boys (presumably leading to the failure to develop age-appropriate assertiveness), whereas coerciveness and threats of rejection are more predictive for girls (presumably leading to low self-concept). For both boys and girls, peer rejection, lack of friends, and engagement in antipathetic relationships (e.g., enemies) in the peer group place children at risk for victimization. Again, it should be noted that these peer-group risk factors are also consequences of victimization, suggesting the vicious cycle between peer victimization and poor peer relations in which children can become trapped.

Although aggressive-victims often have risk factors similar both to aggressors and to victims, there is also evidence of distinct risk factors. In the home context, rates of parental abuse and physical punishment are dramatically higher for aggressive-victims than for other children, and aggressive-victims tend to be rejected more and have fewer friends than either aggressors or victims. Although this entry has not focused on biological origins, it is worth noting that aggressive-victims have high rates of neurological deficits and attention deficit hyperactivity disorder (ADHD) as well. It is believed that these home and peer-group experiences (and possibly the biological risk factors) lead to hostile attribution biases (i.e.,

tendencies to interpret ambiguous behavior by others as hostile in intent), which contribute to aggressive-victims' behavior and further maltreatment by peers.

Subtypes

Despite the general findings regarding aggression and victimization reviewed above, not all acts of aggression are the same. Instead, aggressive behavior can be distinguished in terms of the form that it takes, the function that it serves, and the relationship context in which it occurs.

Forms

Historically, attention has been directed primarily toward studying overt forms of aggression, such as hitting, pushing, or teasing. More recently, however, researchers have realized that aggression also occurs in a more covert form. This type of aggression, variously called relational, social, or covert aggression, includes behaviors such as gossiping, spreading rumors, excluding the victim from groups, and manipulating relationships in a hurtful manner.

Overt versus relational forms of aggression and victimization differ according to age, sex, and context. Developmentally, physical aggression occurs most commonly during early childhood, and verbal forms emerge with increasing language capacities during early to middle childhood; in contrast, relational forms of aggression become more common during adolescence as knowledge of social structure, time spent with peers, and importance placed on peer relations all increase. The historical focus on overt forms of aggression has led to the notion that boys are more aggressive than girls, but more recent considerations of the various forms of aggression have shown that girls and boys are approximately equal in the amount of relational aggression enacted and in the amount received. Finally, there is evidence that different contexts support different forms of aggression, with overt aggression being more commonly enacted on playgrounds and similar areas without adult supervision, whereas relational forms, which might be more difficult for adults to detect, occur more commonly in classrooms.

Despite these differences, there exist high correlations between the two forms of aggression and the two forms of victimization—children who enact high levels of one form tend to also enact high levels of the other, and children who are the victims of one form tend to also be the victims of the other. This has made it difficult for researchers to detect distinct antecedents or consequences of the different forms of aggression or victimization. In other words, the empirical evidence does not provide a clear picture of whether overt versus relational forms of aggression have distinct origins or outcomes, nor whether the victims of these two forms have distinct risk factors or consequences.

Functions

Aggressive behavior can also be distinguished according to the function it serves. Most distinctions by function separate instrumental aggression from reactive aggression. Instrumental aggression (also called proactive aggression) is that which is intended to obtain resources or social status; for example, a child who pushes a peer in order to take a toy. Reactive aggression (also called defensive aggression) is a response, often in an angry, emotionally dysregulated manner, to a perceived offense or threat; for example, the child who throws a temper tantrum and hits a peer during a dispute.

There are two reasons that this distinction according to form is important. First, the two functions of aggressive behavior are believed to have distinct social-cognitive underpinnings. Instrumental aggression is believed (and there is empirical evidence to support this) to be driven by biases in the behaviors considered and evaluations of aggressive behaviors; for example, instrumentally aggressive children tend to believe that positive outcomes will result from aggression and value these outcomes obtained through aggressive behavior. Reactive aggression, on the other hand, is supported by biases in encoding and interpreting social information; for instance, reactively aggressive children tend to interpret others' ambiguous behavior as hostile. A second reason this functional distinction is important is because instrumental aggression and reactive aggression are differentially related to maladjustment. Although both are associated with delinquent behavior, reactive aggression is more strongly related than instrumental aggression to internalizing problems (e.g., depression, anxiety), ADHD symptoms, low prosocial behaviors, and low peer status. It is worth noting that aggressive-only children more often enact instrumental aggression, whereas aggressive-victims more often enact reactive aggression, although the overlap between subgroup classification and functions of aggression is far from complete.

Although the distinction between instrumental and reactive aggression has been important in the study of aggressive children, there has been little attention to how the function of the aggression affects the victims. It seems plausible that distinct characteristics might place children at risk for victimization via instrumental versus reactive aggression, but these distinct risk factors have not been identified. It is also unclear whether victimization by instrumental versus reactive aggression predicts greater maladjustment.

Relationship Contexts

Although researchers have typically considered the characteristics of aggressors and victims in isolation, there is an increasing awareness that aggression often occurs within specific aggressor–victim dyads (i.e., pairs in which a specific child aggresses against another specific child; for instance, Adam aggressing against Billy). For example, one group of researchers observed boys in small play groups and found that more than 50% of aggressive incidents occurred within just 20% of the dyads. These researchers also found that dyads in which aggression was evident on one day tended to be the same dyads that contained aggression on subsequent days. Together, these findings suggest the existence of aggressor–victim dyads in which aggression is especially frequent and persistent across time. The implication of this and related research is that a better understanding of aggression and victimization might be gained by considering the specific dyadic relationships of aggressors and victims.

Although there is very little research adapting this dyadic approach, the limited results demonstrate the importance of considering this relationship context. For instance, it has been found that aggression occurred more commonly within relationships based on mutual disliking (i.e., antipathetic relationships) than within friendships or acquaintanceships with neutral peers. Moreover, victimization within antipathetic relationships was more strongly predictive of maladjustment than was victimization within other relationships, suggesting that victimization within certain relationship contexts (i.e., antipathetic relationships) is more hurtful than victimization within other relationship contexts.

This focus on aggressor–victim relationships is relatively understudied, but it represents a fruitful approach for future research and consideration of occurrences of aggression. Several questions arise from such a consideration: Is there a differential in personal (e.g., physical strength) or social (e.g., popularity) power in aggressor–victim relationships, and does the amount of this power differential predict the form of aggression or outcomes for the aggressors or victims? To what extent are aggressor–victim relationships unidirectional or bidirectional in the enactment of aggression, and what characteristics of the individuals and relationship predict this directionality? Are aggressor–victim relationships relatively stable (i.e., the same aggressors targeting the same victims) or unstable (i.e., aggressors targeting different victims) over time, and what predicts this stability or instability? These questions represent just some that can be asked regarding aggressor–victim relationships; considering and answering these (and similar) questions represent important future directions for research and considerations for those working in applied settings.

Future Directions

Aggression is a common phenomenon in children's lives, resulting in serious maladjustment for both aggressors and victims. The evidence for this statement is conclusive, and the dismissal that such behaviors are just "kids being kids" is incorrect and arguably irresponsible.

Fortunately, research has identified several risk factors for aggression and victimization, providing a point of prevention or intervention for these problems. Unfortunately, translation of this research into application has been rather slow, and the existing intervention efforts have not proven as effective as would be desired. The refinement and widespread implementation of effective prevention and intervention of aggressive behavior represents an important task of educational psychologists (indeed, all professionals working with or studying children).

Part of the difficulty in developing effective interventions may be that for too long, aggression has been viewed as a homogeneous construct. Recent work has identified distinct forms and functions of aggression, as well as expanded consideration of the problem to one of an aggressor–victim relationship. Each of these approaches offers promise in better understanding and treating aggressors and victims.

Recognition and understanding of aggression and victimization holds much promise for reducing these problems. In schools where teachers are aware of school policies on aggressive behavior and have received training to deal with these problems, students tend to view

teachers as more approachable and willing to take action and, more importantly, experience lower rates of aggression and peer victimization. In other words, the first, and perhaps most important, step is simply in recognizing the problem and resolving to do something about it.

Noel A. Card and Abha S. Rao

See also Friendship; Peer Influences; School Violence and Disruption; Self-Esteem; Vicarious Reinforcement

Further Readings

Card, N. A., & Hodges, E. V. E. (in press). Victimization within mutually antipathetic peer relationships. *Social Development.*

Card, N. A., & Little, T. D. (2006). Proactive and reactive aggression in childhood and adolescence: A meta-analysis of differential relations with psychosocial adjustment. *International Journal of Behavioral Development, 30,* 466–480.

Hodges, E. V. E., Card, N. A., & Isaacs, J. (2003). Learning of aggression in the home and the peer group. In W. Heitmeyer & J. Hagan (Eds.), *International handbook of violence research* (pp. 495–509). Boston: Kluwer.

Hodges, E. V. E., & Perry, D. G. (1999). Personal and interpersonal consequences of victimization by peers. *Journal of Personality and Social Psychology, 76,* 677–685.

Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review, 100,* 674–701.

Perry, D. G., Hodges, E. V. E., & Egan, S. K. (2001). Determinants of chronic victimization by peers: A review and a new model of family influence. In J. Juvonen & S. Graham (Eds.), *Peer harassment in school: The plight of the vulnerable and victimized* (pp. 73–104). New York: Guilford Press.

Salmivalli, C. (2001). Group view on victimization: Empirical findings and their implications. In J. Juvonen & S. Graham (Eds.), *Peer harassment in school: The plight of the vulnerable and victimized* (pp. 398–419). New York: Guilford Press.

Schwartz, D. (2000). Subtypes of victims and aggressors in children's peer groups. *Journal of Abnormal Child Psychology, 28,* 181–192.

Smith, P. K., Morita, Y., Junger-Tas, J., Olweus, D., Catalano, R. F., & Slee, P. (Eds.). (1998). *The nature of school bullying: A cross-national perspective.* New York: Routledge.

ALTERNATIVE ACADEMIC ASSESSMENT

Alternative academic assessment (AAA) is a class of procedures that are commonly used to assess student progress within the context of the curriculum to inform instruction. AAA is described as *alternative* because the measurement approaches are often used as an alternative to published norm-referenced and criterion-referenced tests. The latter types of published tests dominated the measurement of academic achievement and aptitude for most of the past 100 years. AAA emerged within the past 30 years as a potential supplement or replacement for more traditional procedures and instruments (e.g., Iowa Test of Basic Skills, Wechsler Individualized Achievement Test).

Many AAA procedures can also be described as *performance-based assessments* or *curriculum-based assessments (CBAs)*. These types of assessments typically rely on task demands and responses that are substantially similar to what might be observed during the process of instruction and learning. Production-type responses are more common than selection-type responses. That is, AAA procedures typically score the examinee's performance on the target behavior rather than a more convenient substitute. If reading is the behavior or interest, then the student's performance within the curriculum context is measured directly. For example, the examinee's oral reading rate might be measured while he or she reads aloud from a sample of curriculum and instructional materials. Similar procedures can be used in place of published tests that contain content dependent on curriculum and instructional materials or selection-type responses (e.g., multiple choice) that indirectly measure the target behaviors. Although there is a wide variety of AAAs, including CBA, curriculum-based measurement, CBA for instructional design, criterion-referenced curriculum-based assessment, curriculum-based evaluation, and informal reading inventories, most are characterized by procedures that measure production-type responses and stimulus content that are substantially similar to that of the curriculum and instructional materials. There are several advantages of AAAs, including efficiency of test development, administration, and scoring; the use of local norms; the utility for benchmarking and progress monitoring; and their utility within the problem-solving model.

Types

Curriculum-Based Assessment

Most AAAs fall under the umbrella of CBA. The fundamental characteristics of CBA include that it

(a) derives from, or is substantially similar to, the curriculum; (b) is linked to instruction; and (c) is used primarily to guide curriculum placement and instructional procedures. CBA procedures and instrumentation can be developed to assess skills within either a broad domain or a narrow domain.

CBA is divided into two subgroups: general outcome measures (GOMs) and subskill mastery measures (SMMs). GOMs are used to assess the level and rate of student achievement within a broad range of skills. GOMs are typically used to assess the achievement within the annual curriculum and instruction. A consistent set of procedures and instrumentation are used throughout the academic year. For example, mathematics computation in second grade might be assessed with tasks that span two-digit addition without carrying through four-digit subtraction with borrowing. The stimulus set and task demands are heterogeneous and representative of the annual curriculum. In contrast, SMMs are used to assess the level and rate of student achievement within a narrow range of skills and, usually, within a narrow range of time, which might be defined by an instructional unit. Procedures and instrumentation might change for each instructional unit so as to assess a specific and distinct set of skills that are aligned with the curriculum and instruction. For example, mathematics computation assessments might be narrowed to include only two-digit by two-digit subtraction without borrowing. The stimulus set and task demands are homogeneous and representative of the short-term instructional goal.

The variety of CBA procedures is broad. The following sections review a subset of available procedures. These include curriculum-based measurement, CBA for instructional design, criterion-referenced CBA, curriculum-based evaluation, and informal reading inventories. There are many other CBA procedures, which include end-of-chapter assessments that come along with curriculum materials. The selected procedures are a representative sample of those with the most prominence.

Curriculum-Based Measurement

Curriculum-based measurement (CBM) is a standardized assessment procedure that is used to index the level and rate of academic growth in four basic skills: reading, mathematics, written expression, and spelling. CBM is generally classified as a GOM because it is used to assess student achievement within the annual curriculum. It is defined as standardized procedure and not as a standardized test. This distinction is necessary because the instrumentation is not standard across applications.

CBM was developed by Stanley Deno and colleagues in the late 1970s through the mid-1980s. More than 200 research studies of CBM are published in peer-refereed journals, and technical development continues through the present day. CBM oral reading fluency (CBM-R) is the most prominent of the available procedures. In CBM-R, the student reads aloud for one minute as the administrator (usually the teacher) follows along on another copy of the passage. The administrator notes errors made and tallies up the number of words read correctly. This yields the student's reading rate, which is reported in units of words read correctly per minute. Error rates and accuracy are also reported sometimes.

CBM was designed to serve as an "academic thermometer" to monitor students' growth in four basic skill domains. The procedures are sometimes described as "dynamic indicators of basic skills." CBM is dynamic in that assessment outcomes are useful to evaluate the effects of instruction over the short term. Outcomes are useful to guide either summative or formative decisions. Summative decisions require one-time evaluations of the level of academic achievement. Formative decisions require ongoing evaluations of the level and rate of academic achievement, which are usually evaluated to estimate the effects of instruction. CBM is an indicator in that it provides a general assessment of academic health within each domain. That is, although CBM-R yields outcomes in words read correctly per minute, that level of oral reading fluency is useful to estimate the general reading achievement and predict performance on large-scale assessments (e.g., statewide tests). CBM targets basic skills in that there are procedures in each of the four fundamental basic skill domains (reading, writing, math, and spelling).

Formative assessment can inform intervention and determine the effectiveness of an intervention. Here is an example: Nick, a second-grade student was referred to the school psychologist for low CBM scores in reading. Nick reads 30 words correctly per minute, compared to the expected 60 words per minute by second grade. The school psychologist or teacher implements an intervention of paired reading for half an hour a day. The following week, Nick completes another CBM at 35 words per minute. Now

the discrepancy between Nick's score and the expected score has decreased, signifying that Nick is responding to the intervention. If the discrepancy continues to decrease, Nick will continue with paired reading until he reads at the expected second-grade level. Should the discrepancy remain the same or increase, the intervention will be modified until the desired level of growth is achieved.

CBA for Instructional Design

CBA for instructional design (CBA-ID) was developed to match the student to the appropriate instructional and curriculum level and focuses on individualizing instruction to ensure mastery learning of all students. A good instructional match maximizes student learning and engagement by fitting the student's needs. Research has found that students are most successful when they are able to respond correctly to material 93% to 97% of the time. This is called the instructional level. CBA-ID starts by finding the student's instructional level and then tests for areas of skill deficits.

CBA-ID has four steps. The first is to choose an appropriate passage for the student to read and test for the student's instructional level by asking them to read 20 to 30 randomly chosen words from the passage. If the student gets five or fewer words wrong, then he or she moves on to the next step. If more than five errors are made, then the student is retested in a passage at a lower level. The second step requires the student to read the passage. As the student reads, the examiner records the student's errors. Afterward, the student is asked questions about the passage to assess how well the student comprehends what he or she read. The last two steps of CBA-ID are to use the assessment information to match the student to appropriate reading instruction and curriculum. Progress monitoring continues as the student is instructed and necessary changes are made.

Criterion-Referenced CBA

Criterion-referenced CBA (CR-CBA) is similar to CBA-ID in that it was developed to determine appropriate instructional materials and strategies. The methods differ, however, in that CBA-CR determines an acceptable level of performance by comparing student scores to a locally normed sample of average peers. Students are tested on items from the curriculum in

order of easy to difficult within a reading series. The teacher or examiner constructs reading tests by using 100-word passages from the beginning, middle, and end of the reading series. Students are tested across 3 days on nine passages; three from the beginning, middle, and end each. As the student reads, the examiner records student errors and then calculates the student's accuracy level (percentage of words read correctly across the entire 100-word passage) and rate of reading (the product of accuracy and 60 divided by total seconds it took the student to read the passage). Next, the student is asked six comprehension questions. Finally, the median scores for the 3 days of assessment are summarized. Assessment decisions are made based on the student's performance compared to the mastery criteria.

Curriculum-Based Evaluation

Curriculum-based evaluation (CBE) is used to determine the student's general whereabouts in the curriculum to identify specific areas of skill deficit. Administration begins with survey level assessment to identify if the student's performance meets expected goals. The initial survey level assessment is developed by sampling a broad domain from the curriculum. In the case of reading, for example, broad skills such as decoding and comprehension are tested. If the results show that the student has a deficit, a skill-specific criterion-referenced test is administered. This second test focuses on specific skills such as segmenting words, rhyming, intonation, and other reading-related subskills. Results are used to adapt instruction to the student's needs.

Informal Reading Inventories

Informal reading inventories (IRIs) are similar to CBE in that they identify specific subskill deficits. There is much variance among the increasing number of published IRIs; however, most follow a typical assessment method. First, the student is asked to read from a graded word list. Depending on the student's accuracy, he or she moves up or down a grade level in word lists until the student's instructional level is found. The definition of this level varies among IRIs but usually falls around 90% words read correctly. Second, the student reads a passage from his or her instructional level. As the student reads, the examiner records the student's errors, or *miscues.* This yields

the student's accuracy and reading rate. Next, the miscues are grouped into categories like omissions, substitutions, low fluency words, and repetitions. Attention is also paid to the types of words that are miscued. This allows the examiner to identify specific areas of weakness. For example, a student may struggle with words with a silent *e*. IRIs determine not only which level of curriculum to place students in but also specific areas of weakness.

Advantages

Efficiency

The administration of AAAs is quick (taking as little as 1 minute) and requires few resources. Items can be developed easily by teachers and taken right from the curriculum. Furthermore, AAAs require little training to administer or score. Also, because they are flexible and can be designed to assess specific domains of academics, AAAs allow school psychologists and other educators to test specific assessment questions rather than routinely test across a broad number of academic domains.

Use of Local Norms

Performances on AAA are often interpreted with reference to local normative data. Local norms are developed from samples of student behavior using AAA procedures. Local norms directly represent the school district population, academic goals, and outcomes rather than the performance of students nationwide. They also decrease the likelihood of bias in decision making because they are representative of student age, grade, race, educational background, and socioeconomic status.

Another advantage of local norms is that there is greater overlap between what is taught and what is tested. Districts have the flexibility to design comparative data based on the specific curriculum. This high teaching-testing overlap yields more meaningful data on student progress. CBM-R is the most common alternative academic assessment that is used to create local norms. In many districts, students complete CBM-Rs in the fall, winter, and spring.

Benchmarking

Benchmarks are sometimes used to evaluate student performance with a criterion-referenced interpretation.

Both benchmarks and local norms are used as referents to evaluate whether students are making adequate progress to achieve expected long-term goals. For example, if students are expected to read 60 words correctly per minute in CBM-R by spring of first grade, then by winter, it is likely they should read 45 words correctly per minute. A student who scores significantly lower than the winter benchmark is identified as at-risk. The student is then given additional help so that by spring, he or she achieves the targeted level of performance. Benchmarking provides information that will help in determining which students are at risk and should be monitored more closely. To this end, benchmarking is also a vital component of prevention. It allows educators to identify and fix small problems before they become larger.

Utility Within the Problem-Solving Model

Historically, students were placed in special education based on their scores on large norm-referenced tests (the "test and place" model). An intelligence score below a certain criterion resulted in a label such as "learning disabled." The student was then matched to special education services based on that label, and the problem was viewed as an inherent trait within the student. Beginning in 2001, a shift has occurred away from the traditional model and toward the problem-solving model.

The problem-solving model focuses not on traits of the student but on environmental and situational factors that can be modified to increase student outcomes. The problem-solving model follows these steps: identify the problem, measure the severity of problem, explore possible interventions, implement an intervention, and measure progress in hopes that the student can be successful in the general education setting. The problem-solving model is a circular process, as progress is constantly being monitored and interventions are adjusted accordingly. The problem-solving model does not rely on unproven inferences as does the traditional model. Hypotheses are continually tested and monitored with AAAs. Thus, decisions are evidence-based because student data are considered when creating and modifying interventions.

AAAs provide the necessary tools to implement the problem-solving model, as they efficiently assess student performance and can be repeated often. As the

popularity of the problem-solving model increases, so does that of AAAs.

Utility for Progress Monitoring

Alternative academic assessments not only identify students who are at-risk but also determine if intervention efforts are successful. Because they are efficient and drawn directly from the curriculum, AAAs can be used repeatedly to monitor student progress and response to intervention. Repeated measures provide not only the level of performance but also the student's rate of growth. This allows educators to determine if the student is making adequate improvement or not. If the intervention does not produce the desired rate of growth, then changes are made to the intervention.

Theodore James Christ and Sarah Scullin

See also Assessment; Evaluation; Measurement; No Child Left Behind; Testing

Further Readings

Andrade, H., & Du, Y. (2007). Student responses to criteria-referenced self-assessment. *Assessment & Evaluation in Higher Education*, 32, 159–181.

Brookhart, S. M. (2007). Editorial. *Educational Measurement: Issues and Practice*, 26(2), 1–2.

Vacca, J. J. (2007). Incorporating interests and structure to improve participation of a child with autism in a standardized assessment: A case study analysis. *Focus on Autism and Other Developmental Disabilities*, 22, 51–59.

Worthen, B. R. (1993). Critical issues that will determine the future of alternative assessment. *Phi Delta Kappan*, 74(6), 444–448.

AMERICAN EDUCATIONAL RESEARCH ASSOCIATION

The American Educational Research Association (AERA) is the primary professional organization for educational psychologists. Founded in 1916, it is concerned primarily with encouraging and promoting scholarly work, including research and writing, that is associated with the educational process. Another important role that AERA plays is the dissemination of information related to its mission.

The 25,000 members of the organization are a collection of scholars and practitioners from diverse fields, well represented by the 12 divisions that the organization hosts (see the following box for the names of these divisions).

American Educational Research Association Divisions

Administration, Organization, & Leadership (Division A)
Curriculum Studies (Division B)
Learning & Instruction (Division C)
Measurement & Research Methodology (Division D)
Counseling & Human Development (Division E)

History & Historiography (Division F)
Social Context of Education (Division G)
School Evaluation & Program Development (Division H)
Education in the Professions (Division I)
Postsecondary Education (Division J)
Teaching & Teacher Education (Division K)
Educational Policy & Politics (Division L)

Each of these divisions (and members often belong to more than one) carries on its own activities, including the election of officers, the publication of a newsletter, the granting of awards for outstanding research by graduate students as well as professionals, and a listserv so that members can remain abreast of new developments.

In addition to its 12 divisions, AERA also sponsors special interest groups (SIGs), which are less formal than divisions but serve many of the same purposes. As is apparent by the list (see the following box), SIGs are of a much more specific nature than are divisions. As of 2006–2007, there are 42 SIGs and each SIG has a program of awards as well.

Special Interest Groups (SIGs)

Action Research
Adolescence and Youth Development
Adulthood and Aging
Adult Literacy and Adult Education
Advanced Studies of National Databases
Advanced Technologies for Learning
Arts and Inquiry in the Visual and Performing
  Arts in Education
Arts and Learning
Arts-Based Educational Research
Associates for Research on Private Education
Bilingual Education Research
Biographical and Documentary Research
Brain, Neurosciences, and Education
Business Education & Computer Information System
Career and Technical Education
Career Development
Chaos & Complexity Theories
Charter School Research and Evaluation
Classroom Assessment
Classroom Management
Classroom Observation
Cognition and Assessment
Communication of Research
Comprehensive School Reform
Computer and Internet Applications in Education
Conflict Resolution and Violence Prevention
Confluent Education
Constructivist Theory, Research, and Practice
Cooperative Learning: Theory, Research and Practice
Critical Educators for Social Justice
Critical Examination of Race, Ethnicity, Class and
  Gender in Education
Critical Issues in Curriculum and Cultural Studies
Critical Perspectives on Early Childhood Education
Cultural Historical Research
Democratic Citizenship in Education
Design and Technology
Disability Studies in Education
Districts in Research and Reform
Doctoral Education across the Disciplines
Early Education and Child Development
Ecological and Environmental Education
Education, Health, and Human Services
Educational Change
Educational Enterprises
Educational Statisticians
Education and Philanthropy
Education and Student Development in Cities
Education and the World Wide Web

Faculty Teaching, Evaluation, and Development
Family, School, Community Partnerships
Fiscal Issues, Policy, and Education Finance
Foucault and Education
Hierarchical Linear Modeling
Hispanic Research Issues
Holistic Education
Home Economics Research
Inclusion & Accommodation in Large-Scale
  Assessment
Indigenous Peoples of the Americas
Indigenous Peoples of the Pacific
Informal Learning Environments Research
Instructional Technology
International Studies
Invitational Education
Ivan Illich
John Dewey Society
Language and Social Processes
Large Scale Assessment
Law and Education
Leadership for School Improvement
Leadership for Social Justice
Learning and Teaching in Educational Leadership
Learning Environments
Learning Sciences
Literature
Lives of Teachers
Longitudinal Studies
Marxian Analysis of Society, Schools and Education
Measurement Services
Media, Culture, and Curriculum
Mentorship and Mentoring Practices
Middle-Level Education Research
Mixed Methods Research
Moral Development and Education
Motivation in Education
Multicultural/Multiethnic Education: Research,
  Theory, and Practice
Multiple Intelligences: Theory and Practice
Multiple Linear Regression: The General Linear Model
Music Education
NAEP Studies (formerly Research Using NAEP Data)
Narrative and Research
Organizational Theory
Out-of-School Time
Paulo Freire
Peace Education
Philosophical Studies in Education
Politics of Education

Portfolios and Reflection in Teaching and Teacher
  Education
Postcolonial Studies and Education
Problem-Based Education
Professional Development School Research
Professional Licensure and Certification
Professors of Educational Research
Qualitative Research
Queer Studies
Rasch Measurement
Religion and Education
Research, Education, Information, and School Libraries
Research Focus on Black Education
Research Focus on Education and Sport
Research Focus on Education in the Caribbean
  and Africa
Research in Global Child Advocacy
Research in Mathematics Education
Research in Reading and Literacy
  (formerly Basic Research in Reading and Literacy)
Research in Social Studies Education
Research on Evaluation
Research on Giftedness and Talent
Research on Learning and Instruction in Physical
  Education
Research on Teacher Induction
Research on the Education of Asian Pacific Americans
Research on the Education of Deaf Persons
Research on the Inclusion of Students with
  Disabilities and Limited English Proficient Students
  in Large-Scale Assessment
Research on the Superintendency
Research on Women and Education
Research Use
Rural Education
Safe Schools and Communities
School Choice

School Community, Climate, and Culture
School Effectiveness and School Improvement
School Indicators, Profiles, and Accountability
  (formerly School Indicators and Profiles)
School/University Collaborative Research
Science Teaching and Learning
Second Language Research
Self-Study of Teacher Education Practices
Semiotics in Education
Service-Learning & Experiential Education
Social and Emotional Learning
Society of Professors of Education
Sociology of Education
Special Education Research
Spirituality & Education
State and Regional Educational Research Associations
Stress and Coping in Education
Structural Equation Modeling
Studying and Self-Regulated Learning
Supervision and Instructional Leadership
Survey Research in Education
Systems Thinking in Education
Talent Development of Students Placed at Risk
Teacher as Researcher
Teacher's Work/Teachers' Unions
Teaching Educational Psychology
Teaching History
Technology, Instruction, Cognition, & Learning
Technology as an Agent of Change in Teaching
  and Learning
Test Validity Research and Evaluation (formerly
  Objective Analysis of Qualitative Research Methods)
Tracking and Detracking
Urban Learning, Teaching, and Research
Vocabulary
Workplace Learning
Writing and Literacies

Neil J. Salkind

Web Sites

American Educational Research Association (AERA):
  http://www.aera.net

AMERICAN INDIANS AND ALASKA NATIVES

American Indians and Alaska Natives (AI/AN) are persons descended from the original inhabitants of North,
Central, and South America and the Caribbean. Those who occupy what is now the State of Alaska are referred to as Alaska Natives. The colonization experience of Alaska Natives is originally seeded in Russian occupation, but their experience mirrors that of the tribes in the lower 48 states. As a group, tribal peoples of the Continental United States are referred to herein as First Nations persons. In the United States, there are more than 560 federally recognized nations and an untold number of non–federally recognized groups. These First Nations are culturally distinct and include populations speaking more than 300 discrete languages. The 2000 U.S. Census Bureau reported First Nations

persons of full or partial decent comprise roughly 1.5% of the total U.S. population, accounting for slightly more than 4 million persons. In 1987, Russell Thornton estimated the population had exceeded 72 million in 1492. By 1800, according to Thornton, the number had been reduced by roughly 95% as a result of disease, warfare, and oppression. Systemically, the fields of education, mental and behavioral health, and medicine are products of this historical context. A conscious endeavor to comprehend the First Nations experience is essential to any professional working with this population. Sadly, this process is generally not undertaken during academic professional training. Following is a discussion of the historical implications for First Nations persons in relation to their inter- and multigenerational experience. Cultural resiliencies and treatment implications are considered, as well as best-practice frameworks.

Historical Context

It is widely recognized that Columbus was not the first European to make contact with the Americas. Regardless, this discovery myth persists and permeates presumptions about the First Nations in many fields, including educational psychology. The consequences of Columbus's contact, however, have been significant. One of the earliest outcomes of his arrival was the enslavement of Indigenous inhabitants. European-modeled slavery directly contributed to the mass disruption of many tribes' gender role structures and systems of government. Tribes were impacted differently by slavery depending upon the era in which they interacted with the newcomers; however, contact generally magnified intertribal disputes and fostered a divide-and-conquer stance toward the First Nations. Perhaps at the height of irony, some First Nations even adopted a pseudo-European model of slavery after years of intermingling with those of European and African descent. When slavery was abolished in the United States, all but one slave-holding tribe extended full tribal citizenship to newly freed slaves. This entitled former African slaves the right to acquire land, tribal representation, and protection under tribal laws, therefore requiring the U.S. government to deal with such persons as African Indians. Slavery historically correlates with the introduction of blood quantum as a measure of ethnic identity and belongingness for First Nations persons, a concept formalized within the federal reservation system. The impact of European-modeled slavery in the Americas is complex and ongoing, and its understanding changes the completion of what has traditionally been perceived as a Black and White element of American history.

More devastating than slavery to the First Nations and their ways of living was the impact of disease post-1492. Of the many diseases brought to the Americas by Europeans, the greatest killer was smallpox. This disease followed trade and warfare routes and struck in repeated waves of pandemic, decimating the vulnerable immune systems of the First Nations. Thornton has identified scores of other diseases introduced by Europeans, including measles, the bubonic plague, cholera, several variations of venereal disease, and rare forms of influenza and respiratory disease. Likewise, the introduction of African slaves to the Americas saw the presentation of diseases such as malaria and yellow fever. Thus, the clear conqueror of the First Nations was the repeated exposure to diseases by which they were decimated at a horrific and incomprehensible rate.

As the First Nations population decreased and the number of Europeans increased, the competition for land and resources became a focal point for conflict. Warring European groups, such as the French and English, capitalized on traditional rivalries between some First Nations tribes and forged alliances with opposing tribes. Most alliances were relatively short lived and dissipated once the First Nations were no longer of benefit to their White allies. Tensions soared between Europeans and the First Nations as the United States of America declared its independence. With the cultural integration of European tools of war such as the horse and firearms, conflicts between the two groups became more intense and lethal. The United States adopted a strategy of treaty making and entered into agreements with First Nations tribes, promising to cease hostilities in return for land concessions on the part of First Nations persons. Between the years of 1775 and 1890, hundreds of treaties were signed between the First Nations and the United States, though few were honored by the United States for any meaningful length of time in their original form. Tribes continued to be encroached upon, and armed conflict flourished. The U.S. Bureau of the Census indicated in 1894 that more than 53,500 American Indians were killed in wars between the United States and First Nations tribes. This number is likely in the hundreds of thousands if one includes the numbers who died as a result of Indian against Indian warfare as an outcropping of some tribes' alliances with the U.S. government.

Acts of genocide add to First Nations casualties. Genocide includes acts intended to destroy (partly or wholly) a national, ethnic, racial, or religious group, including killing or causing serious bodily or mental harm, instituting living conditions highly correlated with death, preventing births to the group, and/or forcibly transferring children of one group to another. Both the U.S. government and its citizens committed acts of genocide. Often cited are incidents of germ warfare via infected blankets given to some First Nations; however, it is difficult to determine how many deaths may have occurred in this fashion, if any. Clear examples of genocide against the First Nations are found in incidents such as the hunting down and murder of First Nations persons during raids in the California and Texas territories, where American Indians were commonly viewed as less than human. In addition, scores have died as a result of harsh governmental policies that fostered little chance for sustenance and survival. It is difficult to discern where acts of warfare end and genocide begins. Many of the incidents of the Indian Wars once described as *battles*, such as those at Sand Creek and Wounded Knee, have now come to be known as massacres of First Nations persons instead.

Though resistance continued on a relatively small scale after 1890, that year is generally recognized as the end of the Indian Wars. As the First Nations fell under the control of the United States, relocation and removal were increasingly used to deal with "the Indian problem." Removals persisted for decades following 1890, and nearly every First Nations group was affected by relocation as the United States strove to accommodate its encroaching settlers. These forced moves separated individuals from their families, communities, and traditional lands upon which entire ways of life and worldview systems resided. During marches, tribes often endured harsh treatment and conditions, cutting to the core of the human capacity to make meaning of what was being endured. High rates of mortality were recorded, and historical writings reveal the emergence of modern-day disorders such as *refugee syndrome* and *concentration camp syndrome*, conditions currently recognized as manifestations of posttraumatic symptoms. The First Nations were faced with repeated and persistent stress, trauma, loss, and grief to which they were forced to respond. A core source of resilience and coping was found in the pan-Indigenous value system and worldview. The First Nations turned to their spiritual leaders for

guidance and hope. As the U.S. government sought continually to manage its Indian problem, it resorted to an apartheid approach of diplomacy—the reservation system. Tribes were generally removed from traditional lands and given dominion over a smaller tract. The life on reservations was often appalling, with starvation, violence, and death all too frequent. The changes First Nations persons faced were pervasive and affected their mental, behavioral, and physical health. The impact of this paternalistic treatment by the United States persists, and First Nations persons continue to struggle with the implications.

As American Indians and Alaska Natives fell under the increased control of Whites, acculturation and assimilation pressures mounted. This is particularly true in respect to the education of First Nations youth. The boarding school era is recognized by First Nations scholars and professionals as the most destructive period in U.S. Indian policy. During the 1800s and 1900s, First Nations children were removed from their homes, as early as age 5, and sent to Christian mission and Bureau of Indian Affairs schools for European-oriented education with a focus on assimilation into White culture. Probably the most famous school was Carlisle Indian School in Pennsylvania, the first off-reservation government-sponsored boarding school. The school was established in 1879 by Henry Pratt, a veteran of the American Civil War and the Indian Wars, whose goal was the complete assimilation of the First Nations. Pratt's motto was "Kill the Indian and save the man." This stance toward Indian education continued well into the 1900s, and First Nations youth were trained in domestic and labor tasks via the school's *outing system* that prepared them for their place in White society. Children were not allowed to practice traditional culture and were prevented from speaking traditional languages and wearing traditional hairstyles and clothing. First Nations youth were forced to practice Christianity and forbidden, often in the face of physical threat, to practice their traditional religions. Children were subjected to harsh punishment in the military fashion of the schools' educational philosophy, and many children endured emotional, physical, and sexual abuse. This era saw generations torn from their traditional, holistic ways of learning and knowledge acquisition and reared instead in a militaristic, institutional setting virtually devoid of the caretaker bonds now recognized as fostering healthy attachment and relationships. The full effect of the boarding school era continues to be

examined in respect to both the costs to and the resiliencies drawn upon by First Nations persons.

As failed Indian policies became apparent to the mainstream citizenry, the political tide turned from paternalism to that of fostering self-determination. One failed attempt at this goal was that of Termination. Termination policy was instituted in the early to mid-1900s to defederalize tribes, dissolving their political status as sovereign nations within the United States and thus their trust relationship with the government. The naive intention was to end governmental paternalism, but what was actually instituted was another form of forced assimilation. First Nations persons were subjected to state laws, and tribal lands were converted to private ownership by former tribal members. First Nations persons were forced to own land individually versus communally and were often forced to utilize it for farming, though virtually no provisions were made for helping tribes obtain the needed capital for such an endeavor. Much of the land made its way to White owners when the Indigenous owners were forced to sell it to support themselves and their families. Virtually overnight, First Nations persons in as many as 100 reservations, bands, and rancherias became *not Indian* as defined by mainstream law. Termination policy has resulted in significant identity struggles for many American Indians and Alaska Natives, as one's ethnicity is defined by another, seemingly at whim and on a continuous basis. An additional outcome of Termination was the mass removal of First Nations persons from reservation areas to urban areas with the promise of employment, education, medical care, and improved quality of life. Unfortunately, what many First Nations families found was poverty and an increased sense of marginalization, as they were now separated from their tribal communities.

The 1960s marshaled in an era of societal and political change in the United States. Self-determination for the First Nations emerged as a priority, and the coming decades saw increased emphasis on fostering sovereignty. The Indian Civil Rights Act of 1968 was passed, prohibiting states from assuming jurisdiction over federally recognized tribal peoples and their lands under Public Law 280. The Indian Education Act of 1972 was an initial effort to require specialized training for educators in an effort to produce and fund cultural competency and to stimulate local attention to First Nations issues. The Indian Self-Determination and Education Assistance Act of 1975 encouraged tribes to assume control over federally funded programs and

provided additional funding. The Indian Child Welfare Act of 1978 was a response to the massive removal and institutionalization of First Nations children via foster care, adoption, and detention in juvenile facilities. First Nations children were to be preferentially placed with First Nations families under the jurisdiction of tribal courts. The American Indian Religious Freedom Act of 1978 recognized the right of First Nations persons to practice their religions and required federal entities to adopt policies of noninterference. In 1988 the Indian Gaming Regulatory Act further defined tribal sovereignty. Also in 1988, Section 5203 of the Tribally Controlled Schools Act added to the intent of the Indian Civil Rights Act and fully repudiated Termination policy. The Native American Graves Protection and Repatriation Act of 1990 acknowledged the profound impact that centuries of objectification have wrought upon the First Nations. The remains of First Nations persons and their burial sites were recognized as sacred, and scores of the deceased were released from museums around the world and returned for proper rites among their peoples. Finally, the Indian Arts and Crafts Act of 1990 furthered the de-objectification of First Nations peoples and helped turn the tide of cultural acquisition.

The 21st Century

Today, the U.S. Census Bureau reports that the First Nations are a young population as compared with other ethnicities, with just less than half living on a reservation or federal trust land and a little more than half living in urban areas. First Nations persons are overrepresented within the numbers of negative social and economic indicators of disparity. Though rates vary widely between tribes and geographic regions, the First Nations find themselves with many of the most disparaging statistical measures of societal success. Economically, they lag behind other ethnic populations, having high poverty and unemployment rates and disproportionately low educational opportunities and graduation rates at all levels. Physical health disparities include high rates of diabetes and heart disease. Mental and behavioral health disparities include depression, posttraumatic stress disorder, and alcohol abuse for the First Nations as a whole.

First Nations children and youth receive disproportionately low levels of prenatal care as compared with other populations. They are exposed to alcohol in utero at higher rates and thus suffer rates of fetal alcohol

syndrome disproportionately. As a group, First Nations youth endure disproportionate rates of diabetes, obesity, inner ear infection, cancer, and toxin exposure. The lack of accessible culturally competent health care compounds the negative impact on First Nations health. Suicide and homicide are among the top 10 causes of death for First Nations youth ages 5 through 14 years, and loss and grief follow this young population as a whole, given the historical and persistent struggles with which it is faced. Emerging public health issues for the First Nations include high rates of pregnancy for young women and girls who have insufficient access to prenatal care, escalating rates of sexually transmitted disease (including HIV/AIDS), and an explosion of gang involvement, even in reservation areas.

Health, Resiliency, and the Balance Way

The disparities and impacts of colonization with which many First Nations struggle is best conceptualized through an Indigenous worldview. The situation then becomes historically bound and has been framed by First Nations elders and tradition keepers as transcending time in a spatial fashion and experienced both individually and collectively. Thus, historical events are experienced in real time by individuals and their communities. Healing occurs through communal support and recognition of suffering and ritual interventions. In this way, suffering is acknowledged by the individual and his or her support system, and assessment, diagnosis, and treatment occur via a vehicle that emphasizes both individual and group strategies. First Nations scholars and clinicians advocate for an intervention framework that recognizes the inter- and multigenerational nature of the loss, stress, and trauma to which American Indians and Alaska Natives are exposed. Such a framework is congruent with the First Nations concept of time and healing and draws upon the use of traditional concepts and strategies.

Assessment, Diagnosis, and Treatment: First Nations Values and Ancient Knowledge

Despite the breath of diversity within the First Nations population, pan-cultural values and worldview perspectives exist. One commonly held view is that of time as cyclical and spatial versus linear, as conceived by the mainstream. Emphasis is on process rather than product. Contextual space or environment is often closely tied with experience and thus with healing. The First Nations see all as connected, whether celestial, elemental, mineral, fauna, animal, or human. Relationships are not compartmentalized along blood lines but rather viewed as broad connections that invoke relational roles. For example, siblings and cousins may be seen as equivalent relationships for an individual, and both will be referred to as brothers or sisters. In this spatial worldview, conceptions of wellness and ill health, life and death are grounded in the idea of keeping balance internally and with the world around one's self, and moving from one time-space domain to the next, respectively. Traditional healers utilize place (e.g., sacred locations) and integrate the help of other elements (e.g., plants, animals) to aid in helping the individual rebalance. Given the value placed on process and the cyclical nature of existence, all human experiences are held as important. Events such as dreams, visions, and premonitions are integrated into the healing process and not pathologized in the vein of mainstream psychology. Traditionally, there is a broad acceptance of difference and individual diversity, and thus a strategy of relative noninterference exists. Persons are supported through change individually with a healer and/or communally, and they are encouraged to find their own path to meaning and balance, utilizing their own gifts and strengths in doing so. This idea of individual difference and noninterference is quite divergent from mainstream thinking and is particularly evident in mainstream socially constructed concepts, such as that of gender or sexual orientation/preference. Pan-culturally, First Nations persons traditionally view gender as discrete from sexual orientation, identity, and preference.

Given the historical context within which the fields of psychology and education have developed and the current statistics the First Nations face, it is clear that pervasive cultural competence is lacking in assessment, diagnosis, and treatment of this population. Stereotyping, stigma, and discrimination pervade the intervention process with the First Nations and are highly correlated with low rates of contact and retention within the helping professions. Clinicians must consider the possibility of institutional distrust on the part of First Nations clients. Accommodation must be made for cultural differences present between the client and the clinician as well as between the client and the system of service. Language can pose a particular roadblock to intervention, regardless of whether the client speaks English as a second or a first language. Previous generations pass down the cultural worldview housed in Indigenous languages. First Nations

languages are relational and descriptive in nature and do not accommodate compartmentalization as English does. Most Indigenous languages provide an understanding of the world as either animate or inanimate, not living or dead. Gendered language too is relatively nonexistent in the fashion of Western-mainstream languages. A close examination of popular standardized assessment and diagnostic tools quickly reveals their inability to competently accommodate First Nations clients. The vast majority of such tools, including the *Diagnostic and Statistical Manual of Mental Disorders* (Fourth Edition, Text Revision) *(DSM-IV-TR)*, have been developed through a linear, White mainstream worldview, and few have been standardized or explored in relationship to their use with the First Nations.

Learning: Indigenous Science and Knowledge Acquisition

American Indians and Alaska Natives have had sophisticated systems of hypothesis testing and knowledge acquisition for thousands of years. This knowledge has persisted in oral, written, pictorial, and ritual traditions. Unfortunately, systemic racism, discrimination, and ignorance have all played a role in the perpetuation of the stereotype of First Nations science as a proto-knowledge, a less sophisticated form of Western-mainstream constructs. It is in this atmosphere that First Nations persons are educated by mainstream institutions of learning. Acculturation and assimilation pressures are significant for First Nations children in educational settings and persist through higher education.

Preschoolers may encounter difficulties adjusting to their new setting and its demands. Traditionally, First Nations youth are raised with close attention to attachment building and may share a bed with their primary caretaker(s), may be breastfed until they are 3 or 4 years old, and may enjoy the attention of multiple caretakers regardless of blood ties. Children entering elementary school are often encountering mainstream culture for the first time and can be shaken by the shift in worldviews within which they must function. Boys may be ridiculed for keeping their hair traditionally long, and all youth are subject to defending themselves against the onslaught of holidays and practices celebrated in school systems that may be Christian-focused or U.S. nationalist. Language issues can be a particular challenge, and children may be required to shift from an experientially based traditional educational focus at home to a more linear, prescribed learning style in the educational setting. By middle and high school, First Nations youth often confront the full force of stereotyping and discrimination, as well as the aforementioned risk factors. They may encounter existential crises, struggling to integrate traditional spiritual beliefs with mainstream culture. The risk for internalized oppression is great, as teens strive for identity and self-preservation. A traditional adolescent may attend a school that promotes stereotyping and demeans the spiritual worldview of First Nations persons via Indian mascots or the promotion of, and forced participation in, Christian-based activities. This pressure mounts as those First Nations youth who do graduate from high school attempt to make their way to college. College students may have to travel great distances from their tribal communities and lands, deepening existential struggles. In higher education they find few First Nations mentors, little funding, and can struggle greatly to resolve the rift between their traditional worldview and that of the Western mainstream.

Indigenizing the Mainstream

Best practices with American Indians and Alaska Natives mandate the integration of traditional knowledge, practices, community, and tradition keepers. Successful strategies and programs are individualized and recognize the potential for divergent worldviews and diversity within this group. Legitimate ways of knowing and healing are held within the First Nations culture and have been utilized for thousands of years to educate and heal this group. The educational psychologist will find a wealth of helpers within the community's natural supportive structure. Elders, traditional healers and mentors, extended family, and many others can be of assistance for case conceptualization and treatment. Traditional knowledge and values can be found in a group's original instructions (creation story) and provide a useful framework for conceptualizing a client's struggle, as well as his or her ethnic identity. Finally, policies that support the integration of understanding between the mainstream and First Nations may be supported by educational psychology as socially just objectives that promote healing and understanding for both groups.

Leah M. Rouse Arndt

See also Bilingualism; Cognitive and Cultural Styles; Culture; Discrimination; Ethnicity and Race

Further Readings

Duran, E., & Duran, B. (1995). *Native American postcolonial psychology.* Albany: State University of New York Press.

Duran, E., & Ivey, A. (2006). *Healing the soul wound: Counseling with American Indians and other Native Peoples.* New York: Teachers College Press.

LaFromboise, T., & Dizon, M. (2003). Indigenous American Indian population groups. In J. T. Gibbs (Ed.), *Children of color: Psychological interventions with culturally diverse youth* (2nd ed., pp. 41–91). San Francisco: Jossey-Bass.

Thornton, R. (1987). *American Indian holocaust and survival: A population history since 1492.* Norman: University of Oklahoma Press.

ANDROGYNY

Androgyny is a term derived from the Greek *andras* (άνδρας—man) and *gyne* (γυνή—woman) referring to either the absence of any distinguishing masculine or feminine traits, as in the Hijras of India, or the combination of both masculine and feminine characteristics, whether spiritual, psychological, or physiological.

Most Western cultures presume a binary opposition between male and female. In the 1950s, June Singer revived a mystical interest in androgyny, reconciling the "masculine" and "feminine" aspects of a single human, restoring the balance between what Jung called *animus* and *anima.* Like Mircea Eliade and Carl Jung, Singer treated androgyny as archetypal, in which the divided self yearned for the complete reunion of male and female. This understanding of androgyny as a metaphysical ideal was implicit in shamans or deities like Buddha, Shiva, Kuan Yin, and Elohim. Even so, Singer believed that the sexes were naturally differentiated: that males are generally aggressive, dominant, hard, and logical, and women are passive, compliant, soft, and intuitive.

In 1974, Sandra Bem published the Bem Sex Role Inventory (BSRI), a self-test listing 20 socially desirable female traits, 20 socially desirable male traits, and 20 considered to be neutral. Male traits included "forceful," "analytical," and "self-sufficient"; female traits included "sympathetic," "loyal," and "compassionate"; and neutral items included "truthful," "sincere," and "friendly." Scores revealed the respondent's self-reported possession of socially desirable, stereotypically masculine and feminine personality characteristics. An individual who received high scores for both female and male traits was defined as androgynous, whereas one with low scores in both was described as undifferentiated. Gender traits had little correlation with the ascribed sex of participants. Like Singer and Jung, Bem believed that people who had androgynous psychological traits were the most effective and well-functioning individuals in society.

At this time, Constructivism claimed that gender was socially constructed and could therefore be changed at will. Resistance to gender binaries and heterosexuality took the form of transvestitism, or performances as drag queens or drag kings. Kate Bornstein, having performed as a cross-dressing performance artist and encouraged the self-construction of "who you are," used surgery to change herself in 1998 to a "male-to-female lesbian transsexual" but has now settled into being neither male nor female, a gender outlaw.

To describe a born male as "lesbian" indicates some of the conceptual change required by this new gender fluidity, but it did not necessarily accommodate androgyny. A medical category, gender identity disorder, was created to describe those who felt incompatibility between their felt identity and their anatomy. Improvements in surgical processes made it possible to normalize anatomies as normalized male or female, and medical research sought to explain sex "transgressions" (gender identity disorder, cross-dressing, or homosexuality) physiologically in order to remove blame and effect a "cure." A few transsexuals, like female-to-male Jamison Green, rejected such normalizing "cures" and accepted their androgynous status to the extent of having hormone treatment but not requiring a surgically constructed penis or denying their past.

Despite a relatively low level of sexual dimorphism in humans, Charles Darwin had naturalized the sex binary in *The Descent of Man* by referring to naturally selected sex differences between male and female in gonads, sex organs, body mass, amount and placement of body hair, intelligence, psychological traits such as aggression, and child-rearing practices. But he tended to overlook the high amount of androgyny in the natural world, for instance, in worms and snails. About 30% of the fish species on a coral reef start out as males and end as females, or vice versa, or are both male and female at the same time. Could humans be naturally androgynous?

In 1993 in *Bodies That Matter*, Judith Butler argued that even though we construct gender, our material bodies sometimes prevent us from conforming to social norms. Earlier in *Gender Trouble*, she

claimed that drag queens in their "queer performativity" demonstrate resistance to being required to "perform" normal dichotomous roles of male or female which they do not feel. The existence of physically androgynous humans challenges those dichotomies at an even more profound level.

Anne Fausto-Sterling estimates conservatively that 1 in 1,000 persons is born with androgynous physiological features. Previously called *hermaphrodites*, they are now medically defined as *intersex*, a term, like *androgyny*, applied to any person with characteristics determined as neither exclusively male nor female, or combining features of both. The most common cause of so-called sexual ambiguity is *congenital adrenal hyperplasia*, an endocrine condition in which the adrenal glands produce unusually high levels of virilizing hormones. In genetic females, this leads to an appearance that may be slightly masculinized (large clitoris) to quite masculine. Another form of intersex is *androgen insensitivity syndrome*, in which people born with masculinizing Y chromosomes do not develop male morphology. Other physiological androgynes show chromosome variations such as 47XXY, 45XO, or mosaics. Far from being revered as complete or ideal persons, such anatomical androgynes are usually classified as male or female sex at birth and surgically or hormonally transformed into either male or female in childhood, with remaining variations of the male-female binary seen as transgressive.

Bem and Fausto-Sterling argue that the danger of perpetuating the male-female binary lies in the fact that cultural roles and norms remain dictated by males and highly polarized male values. The placement of the androgyne at the center of our understanding of physical and cultural humanness and our acceptance of complex combinations of male and female defuse current hegemonies.

Felicity Ann Haynes

See also Gender Bias; Gender Differences; Gender Identity

Further Readings

Barbin, H. (1980). *Herculine Barbin, being the recently discovered memoirs of a nineteenth-century hermaphrodite* (M. Foucault, Introduction, & R. McDougall, Trans.). New York: Pantheon. (Original work published 1978)

Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology, 42,* 155–162.

Bem, S. L. (1993). *Lenses of gender: Transforming the debate on sexual inequality.* New Haven, CT: Yale University Press.

Bornstein, K. (1995). *Gender outlaw: Men, women and the rest of us.* New York: Routledge.

Bornstein, K. (1997). *My gender workbook: How to become a real man, a real woman, the real you, or something else entirely.* New York: Routledge.

Butler, J. (1990). *Gender trouble: Feminism and the subversion of identity.* New York: Routledge.

Butler, J. (1993). *Bodies that matter: On the discursive limits of "sex."* New York: Routledge.

Dreger, A. (Ed.). (1999). *Intersex in the age of ethics.* Hagerstown, MD: University Publishing Group.

Dreger, A. (2000). *Hermaphrodites and the medical invention of sex.* Cambridge, MA: Harvard University Press.

Eliade, M. (1965). *The two and the one* (J. M. Cohen, Trans.). Chicago: University of Chicago Press. (Original work published 1962)

Fausto-Sterling, A. (2000). *Sexing the body: Gender politics and the construction of sexuality.* New York: Basic Books.

Green, J. (2004). *Becoming a visible man.* Nashville, TN: Vanderbilt University Press.

Haynes F., & McKenna, T. (Eds.). (2001). *Unseen genders: Beyond the binaries.* New York: Peter Lang.

Singer, J. (1989). *Androgyny: The opposites within* (2nd ed.). Boston: Sigo Press.

ANXIETY

Anxiety is a common mental health concern facing many Americans today. In 1997 Thomas Huberty defined *anxiety* as a unique emotional state characterized by feelings of distress and tension about real or anticipated threats that may manifest in cognitive, behavioral, or physiological patterns. Anxiety can have devastating effects on individuals, as it can interfere with their learning and social and emotional development. In this entry, general information about anxiety is presented. Common features found among the anxiety disorders, types of anxiety disorders, and etiological factors underlying anxiety disorders are discussed. Prevention and intervention strategies are covered.

Components of Anxiety

Anxiety is a complex emotional state and may involve and influence multiple domains of an individual's functioning. Specifically, an individual may experience

cognitive, behavioral, and physiological effects. Common cognitive symptoms of anxiety include excessive worries, concentration difficulties, and memory and attention problems. Anxiety may also be manifested through such behavioral symptoms as motor restlessness, difficulty sitting still, and attempts to escape or avoid anxiety-provoking stimuli or situations. Finally, anxiety also includes physiological symptoms, such as muscle tension, increased perspiration, rapid heartbeat, headaches, and stomachaches.

Anxiety as a Unique Emotion

Anxiety is a unique emotion as it can be viewed in both a positive or negative light. A slight amount of anxiety can be helpful and facilitate an individual's performance, whereas too much anxiety can be debilitating and hinder one's performance. In small amounts, anxiety can serve as a motivator and lead to optimal performance in school, work, sports, or other areas in an individual's life. For example, a student can become slightly anxious before a major exam. The slight anxiety felt can motivate the student to study for the exam and do better because of the time spent preparing for the exam. In contrast, high levels of anxiety may interfere with the student's ability to concentrate, process information, or retrieve information from long-term memory. Under these circumstances, the student is less likely to perform his or her best on the exam.

Anxiety can also alert an individual to a potential danger. The fight-or-flight response, also referred to as the acute stress response, involves the activation of the sympathetic nervous system in an emergency situation. The individual will respond to a threatening or dangerous situation by fighting or fleeing. Thus, many believe that anxiety serves as a survival mechanism and protects the individual from harm.

Besides its positive and negative aspects, anxiety can be viewed as a normal indicator of development. During the normal course of development, individuals experience fears and anxieties, but the specific fears and anxieties experienced vary as a function of age. Moreover, there is some evidence to suggest that the number of specific fears and anxieties decrease with age, whereas others suggest that the number of specific fears and anxieties remain the same across the life span. Sources of anxiety for infants include loud noises, strangers, and novel stimuli, and for toddlers, separation from major attachment figures and imaginary creatures. Children fear large animals, darkness,

and natural events, and adolescents fear social alienation. Sources of anxiety for adults include natural events, injury, and financial issues. Most individuals experience these age-specific anxieties and fears, which are mild and transient in nature.

Prevalence of Anxiety and Comorbid Conditions

Prevalence rates of anxiety in community samples are difficult to estimate, especially given the fact that internalizing disorders, such as anxiety, are often difficult to observe and identify. Prevalence estimates for clinical levels of anxiety in children, young and middle-age adults, and older adults range from 5% to 19%, 6% to 8%, and 9% to 11%, respectively. In general, the prevalence of anxiety disorders has been found to increase with age in the child and adolescent population, decline during the young and middle-age adult years, and slightly increase in the older adult years. Gender differences have also been found in the literature. Specifically, females typically report more anxiety symptoms than do males. However, it remains unknown at the present time whether this gender difference is due to females actually experiencing more anxiety symptoms than males or whether females are simply better able to recognize and report their symptoms of anxiety than are males. On the other hand, overall, more females than males are believed to suffer from an anxiety disorder. However, the gender ratios differ based on the type of anxiety disorder diagnosed and the age of the individual. Despite this variability and the need for further research with regard to prevalence rates among different ages and genders, it is clear that anxiety continues to be a major problem for individuals of all ages and one that can potentially lead to significant difficulties within multiple domains of functioning.

Anxiety disorders have high rates of comorbidity with other disorders. The rate of comorbidity between anxiety and depressive disorders may be as high as 55% to 65%. Speculation as to why these rates are so high is that both anxiety and depression share a similar trait known as negative affectivity, or emotional distress. Negative affectivity includes such affective states as worry, self-dissatisfaction, and sadness. High comorbidity rates may also be due to a sequential link between anxiety and depression, with anxiety serving as an early precursor to a depressive disorder. Comorbidity rates are also high between different types of

anxiety disorders. It is not uncommon for an individual who has one anxiety disorder to be diagnosed with another anxiety disorder. Other common comorbid conditions include substance use disorders and disruptive behavior disorders, such as attention deficit hyperactivity disorder, oppositional defiant disorder, and conduct disorder.

From a trait model perspective, anxiety is viewed as a stable personality characteristic. Without treatment, anxiety disorders may persist. Approximately 45% to 65% of individuals diagnosed with an anxiety disorder do not show remission of symptoms. However, approximately 35% to 55% of individuals do show remission, but many who show remission develop other disorders, especially other anxiety disorders. Anxiety disorders interfere with the social and emotional well-being of individuals. If individuals are still in school, academic development may be impaired. For adults, work productivity may decline and unemployment may increase.

Developmental Precursors or Etiological Factors

Different theories exist about the development of an anxiety disorder. The three most popular models used to explain the development of an anxiety disorder are the biological, behavioral, and cognitive models. Biological explanations of anxiety focus on genetics, neurotransmitters, differences in structural regions of the brain, abnormalities in the immune system, and behavioral inhibition. Genetics is believed to play a role in the development of an anxiety disorder. Genetic influences account for approximately 30% to 35% of the variance in anxiety in most cases, suggesting that anxiety is moderately inheritable. The neurotransmitter gamma aminobutyric acid (GABA) has received some attention as a possible risk factor in the development of an anxiety disorder. GABA may increase excitatory responses to real or perceived threats, or it may fail to send messages to inhibit these responses. Perturbations in the hypothalamic-pituitary-adrenal axis indicate structural brain differences in individuals with an anxiety disorder. Behavioral inhibition may be another possible etiological factor. Behavioral inhibition characterizes a child's temperament. Children with this type of temperament are shy and exhibit inhibited behaviors in response to novel stimuli. These individuals are also highly physiologically reactive to such stimuli.

Behavioral explanations for the development of an anxiety disorder focus on learned behaviors. According to behaviorists, anxiety is a learned behavior that is acquired and maintained through a combination of classical and operant conditioning, operant conditioning alone, or modeling. From a classical and operant conditioning perspective, anxiety problems result when a neutral stimulus, such as a large dog, is paired repeatedly with an aversive stimulus (i.e., an unconditioned stimulus), such as a loud noise, to produce an unconditioned response, such as a startled reaction. Through repeated pairings with the unconditioned stimulus, the neutral stimulus becomes a conditioned stimulus, and this conditioned stimulus will produce a conditioned response. In other words, the presence of a large dog will result in a startled response without the loud noise being present on a regular basis. The conditioned stimulus (i.e., a large dog) will then be avoided, and by avoiding the conditioned stimulus, an individual's anxiety is reduced. The avoidance behavior demonstrated by the individual in response to the large dog is an example of operant conditioning. In operant conditioning, the stimulus, task, or situation feared is maintained by a negative reinforcement contingency. The feared stimulus, task, or situation is avoided, and the avoidant behavior is maintained because it reduces the individual's anxiety. In modeling, the individual observes the behavior of significant others in response to aversive stimuli, tasks, or situations. When significant others exhibit avoidance behavior and anxiety in response to aversive stimuli, the individual learns to model these behaviors. Exposure to similar aversive stimuli, tasks, or situations will produce similar behaviors.

The cognitive approach to anxiety disorders assumes that distorted cognitions are responsible for symptom manifestation. According to cognitive psychologists, individuals who have an anxiety disorder or experience high levels of anxiety exhibit threat-related attentional and interpretative biases. These individuals attend to threat-related stimuli, and they interpret ambiguous or neutral stimuli as threatening.

Common Features of Anxiety Disorders

Although different types of anxiety disorders exist, according to Michael Telch, Jasper Smits, Matt Brown, and Victoria Beckner, there are some common features across these different disorders. Common features

include escape and avoidance behaviors, chronic worry, attentional hypervigilance, faulty threat perception, and sympathetic activation. Individuals with an anxiety disorder try to avoid or escape from stimuli or situations that make them anxious, and they constantly worry about current and future events. These individuals attend excessively to cues that they perceive as threatening. The excessive attention given to these cues is referred to as *attentional hypervigilance*. Faulty threat perception is another common feature found among individuals with an anxiety disorder. These individuals erroneously perceive situations as threatening. Sympathetic activation is also a core feature found among individuals with an anxiety disorder. Activation of the sympathetic nervous system producing physiological changes in the body occurs in individuals with an anxiety disorder when there is no real or potential threat. Physiological changes experienced by these individuals in the absence of a real or potential threat may include accelerated heart rate, muscle tension, and increased perspiration and respiration.

Types of Anxiety Disorders

The *Diagnostic and Statistical Manual of Mental Disorders* (Fourth Edition, Text Revision) *(DSM-IV-TR)* has identified 15 types of anxiety disorders, including generalized anxiety disorder, separation anxiety disorder, specific phobia, social anxiety disorder, obsessive-compulsive disorder, posttraumatic stress disorder, acute stress disorder, panic attack, panic disorder with and without agoraphobia, agoraphobia without a history of panic disorder, anxiety disorder not otherwise specified, anxiety disorder due to a general medical condition, and substance-induced anxiety disorder. Generalized anxiety disorders, phobias, obsessive-compulsive disorders, panic disorders, and posttraumatic stress disorders are the most common types of anxiety disorders.

Although there are several types of anxiety disorders, each involves an excessive degree of worry or fear about certain stimuli, situations, or events, which significantly interferes with an individual's normal state of functioning. Generalized anxiety disorder is characterized by a chronic, excessive, and uncontrollable degree of worry about a variety of events or situations, such as friends, family, school, work, or the future. Symptoms of generalized anxiety disorder include fatigue, irritability, restlessness, muscle tension, and difficulties with concentration and sleep. Another common anxiety disorder, specific phobia, is characterized by an extreme

and irrational fear in response to a specific stimulus, such as animals or insects (e.g., dogs), aspects of the natural environment (e.g., storms), blood (e.g., viewing blood or receiving an injection), situations (e.g., being in small spaces), or other stimuli (e.g., loud sounds or costumed characters). This worry must be present for at least 6 months and may lead to symptoms in children such as crying, clinging, tantrumming, dizziness, shortness of breath, and fainting.

Like a specific phobia, social phobia is associated with particular circumstances and must involve symptoms present for at least 6 months. An individual with a social phobia experiences extreme worry regarding social situations. The individual may worry over or fear the possibility of ridicule, humiliation, or embarrassment in social situations, such as speaking in class or conversing with peers. Individuals with a social anxiety disorder may attempt to avoid or escape social behaviors, and they often have poor social skills. These individuals may also experience symptoms such as trembling hands or voice, perspiration, muscle tension, and blushing.

Obsessive-compulsive disorder involves obsessions (recurrent or persistent thoughts or worries that intrude on, and interfere with, an individual's normal functioning) and compulsions (repetitive behaviors, rituals, or practices in which the individual engages to provide relief from, or comply with, the obsessive thoughts or worries). Common obsessions and compulsions in children include contamination (hand washing), safety (checking), preoccupations with orderliness and symmetry (ordering, aligning), and counting or touching rituals. Individuals with an obsessive-compulsive disorder may feel embarrassed by their compulsions and may experience difficulties related to concentration, preoccupations, and perfectionist tendencies.

Posttraumatic stress disorder refers to stress or worry experienced by an individual following a traumatic event (such as a serious injury, death, or catastrophic event). The individual reexperiences the event (via flashbacks, nightmares, or images), as well as the accompanying physiological arousal, and may attempt to avoid stimuli associated with the event. Following the event, the individual may feel helpless, fearful, agitated, or disorganized and may experience hypervigilance, irritability, and difficulties with concentration or sleep. These symptoms must be present for at least 1 month following the traumatic event. If symptoms are present for less than 1 month, then an acute stress disorder may be present. Similar to posttraumatic stress

disorder, acute stress disorder also results from witnessing or experiencing a traumatic event. Symptoms include reexperiencing the event and the accompanying physiological arousal; however, these symptoms are present for no more than 1 month.

A panic disorder refers to recurrent and unexpected panic attacks that are followed for at least 1 month by concern about, or consequences of, having another attack and/or a change of behavior related to the attack. Panic attacks develop abruptly, often last for approximately 10 minutes, and involve symptoms such as heart palpitations, sweating, chest pain, dizziness, fear of dying, feelings of detaching from one's body, and feelings of losing control. If the attacks become more frequent, the individual may come to fear experiencing a panic attack in public places, and develop agoraphobia (fear of public places) as well. However, agoraphobia may also develop in the absence of, and without resulting from, a panic disorder. In these cases, individuals may avoid public situations, such as being in a crowd or traveling in a train, and may experience symptoms of panic in these situations.

Separation anxiety disorder is an anxiety disorder commonly found in children. Separation anxiety disorder refers to excessive and unrealistic worry in response to separation from home or a caregiver. Children with separation anxiety disorder may experience nightmares with separation themes, headaches, stomachaches, or nausea. These symptoms must be present for at least 4 weeks. Separation anxiety disorders tend to decrease with an increase in age. That is, this type of anxiety disorder is common during the childhood years but declines during the adolescent and adulthood years.

Anxiety disorders may also result from external factors, such as a general medical condition or substance use. Finally, for individuals who experience symptoms of anxiety, but whose symptoms, duration, or impairment do not meet the criteria for a specific disorder, a diagnosis of anxiety disorder not otherwise specified (NOS) may be appropriate.

Anxiety Disorders in Children and Adolescents in School

Many children and adolescents who either have been diagnosed with an anxiety disorder or have high levels of anxiety experience difficulty in the school setting. Students with anxiety disorders or high levels of anxiety have more difficulty learning new material,

receive poorer grades, and do not perform as well on standardized and classroom tests. These students may struggle in core courses such as reading and math. They are also more likely to repeat a grade and drop out of school.

Besides experiencing academic difficulties, children and adolescents with anxiety disorders experience poor peer relationships. Because of cognitive distortions or maladaptive thinking patterns, these children and adolescents view their relationships with others more negatively. These negative perceptions of their relationships with others reduce the likelihood of interactions with peers. These individuals may feel socially isolated and experience depression and feelings of hopelessness.

Symptoms of anxiety can significantly interfere with children or adolescents' social-emotional and academic functioning. In light of these concerns, children and adolescents with an anxiety disorder may be eligible for special education and related services under the Individuals with Disabilities Education Improvement Act of 2004 (IDEA). Specifically, anxiety disorders are categorized under the emotional disturbance (ED) category of disabilities. To meet the criteria for an emotional disturbance, a student must exhibit one or more of the following conditions, and the condition(s) must have occurred over a long period of time and to a marked degree and must adversely affect the individual's educational performance:

1. An inability to learn that cannot be explained by intellectual, sensory, or health factors

2. An inability to build or maintain satisfactory interpersonal relationships with peers and teachers

3. Inappropriate types of behavior or feelings under normal circumstances

4. A general pervasive mood of unhappiness or depression

5. A tendency to develop physical symptoms or fears associated with personal or school problems (34 C.F.R. § 300.8)

Assessment of Individuals With an Anxiety Disorder

To detect an anxiety disorder or high levels of anxiety in individuals of different ages, a multimethod approach

to the assessment of anxiety is recommended. A multi-method approach involves the use of different types of measures completed by multiple informants across multiple settings to detect anxiety and comorbid conditions. A variety of assessment techniques, including clinical interviews, rating scales, direct observations, self-report, and psychophysiological measures, are available to assess anxiety in individuals of different ages. Many of these measures are completed by multiple informants (self, parent, spouse, and/or teacher) across multiple settings (home, school, and/or work).

Treatment of Anxiety

Once an assessment or evaluation is completed and high levels of anxiety are detected or an anxiety disorder is diagnosed, assessment results are linked to interventions to ameliorate anxiety and its negative effects. Different treatment strategies are available to address anxiety, including pharmacotherapy, behavioral strategies, and cognitive-behavioral interventions. Additional strategies may also be used to address comorbid issues. Thus, a multimodal approach, consisting of two or more interventions, is often used to alleviate an individual's anxiety and its negative effects.

Pharmacological treatment is one means of alleviating anxiety in individuals. Medications that have been used to treat anxiety include benzodiazepines, selective serotonin reuptake inhibitors, tricyclic antidepressants, and buspirone. Medication is often used in combination with other treatments such as cognitive-behavior therapy because although the medication may reduce anxiety symptoms, it does not help individuals learn to cope effectively with their anxiety.

Behavioral interventions are another means of reducing anxiety in individuals. Relaxation training, systematic desensitization (graduated exposure), and modeling are some of the behavioral strategies used to treat anxiety. These strategies have been shown to be effective. Relaxation training may include deep breathing exercises or progressive muscle relaxation. Progressive muscle relaxation involves individuals learning to relax and tense different muscle groups in order to become more relaxed. Relaxation training may also be found in systematic desensitization. In systematic desensitization, a fear hierarchy is created, consisting typically of 10 to 15 steps evenly spaced. For example, if an individual had a fear of large dogs, the first step in the fear hierarchy may consist of a discussion about dogs. The second step may involve

looking at a picture of dogs. The third step of the fear hierarchy may involve driving past a pet shop and so on until the last step, when the individual pets a real-life dog. The purpose of creating a fear hierarchy is to gradually expose the individual, step by step, to the feared stimulus. The graduated exposure can be conducted using imagery or real-life experiences. Relaxation or another incompatible response to anxiety is induced along the way to calm the individual as graduated exposure of the feared stimulus occurs. Modeling is another behavioral strategy used to reduce fears and anxieties. Modeling is based on social learning theory in which an individual observes, either live or on film, a person who interacts successfully with the feared stimulus or situation. The model is typically of the same age and gender as the individual. After watching the model interact successfully with the feared stimulus or situation, the individual is more likely to perform the same behavior, and the fear and anxiety associated with the feared stimulus or situation are reduced.

Cognitive-behavioral strategies, such as self-instruction, self-control training, and rational-emotive therapy, have also been used to alleviate individuals' anxieties. Self-instruction involves the use of positive self-talk to handle anxiety-provoking situations. In self-control training, individuals learn to modify and restructure maladaptive thoughts, resulting in less anxiety in the presence of anxiety-provoking stimuli or situations. Less anxiety experienced then leads to positive changes in behavior because these individuals are more likely to approach the feared stimuli or situations. Replacement of false, irrational beliefs that underlie an anxiety problem with rational beliefs is the focus of rational-emotive therapy.

Prevention of Anxiety

Because anxiety is a common mental health concern facing many Americans today, efforts should be directed toward the prevention of anxiety disorders. The emotional, social, and economic costs associated with anxiety disorders are astronomical. Economic costs alone are estimated to be more than $40 billion per year. Yet, few prevention programs exist. Although prevention programs are costly up front, universal (primary), selective (secondary), and indicated (advanced) prevention programs are needed. Future efforts should be directed toward the development and implementation of these programs, as there will never be enough

mental health professionals to provide adequate treatment of anxiety and other disorders.

Patricia A. Lowe and Jennifer M. Raad

See also *Diagnostic and Statistical Manual of Mental Disorders*; Personality Tests; Psychoanalytic Theory

Further Readings

Albano, A. M., Causey, D., & Carter, B. D. (2001). Fear and anxiety in children. In C. E. Walker & M. C. Roberts (Eds.), *Handbook of clinical child psychology* (pp. 291–316). New York: Wiley.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., Text rev.). Washington, DC: Author.

Chorpita, B. F., & Southam-Gerow, M. A. (2006). Fears and anxieties. In E. J. Mash & R. A. Barkley (Eds.), *Treatment of childhood disorders* (3rd ed., pp. 271–335). New York: Guilford Press.

Compas, B. E., & Oppedisano, G. (2000). Mixed anxiety/ depression in childhood and Adolescence. In A. J. Sameroff, M. Lewis, & S. M. Miller (Eds.), *Handbook of developmental psychopathology* (2nd ed., pp. 531–548). New York: Kluwer Academic/Plenum.

Dozois, D. J. A., & Dobson, K. S. (Eds.). (2004). *The prevention of anxiety and depression: Theory, research, and practice.* Washington, DC: American Psychological Association.

Evans, D. L., Foa, E. B., Gur, R. E., Hendin, H., O'Brien, C. P., Seligman, M. E. P., et al. (2005). *Treating and preventing adolescent mental health disorders: What we know and what we don't know.* New York: Oxford University Press.

Foa, E. B., & Andrews, L. W. (2006). *If your adolescent has an anxiety disorder.* New York: Oxford University Press.

Huberty, T. J. (1997). Anxiety. In G. Bear, K. Minke, & A. Thomas (Eds.), *Children's needs II* (pp. 305–314). Bethesda, MD: National Association of School Psychologists.

Individuals with Disabilities Education Improvement Act of 2004, P.L. No. 108–446, 70 Fed. Reg. 118 (2004). Retrieved June 28, 2005, from http://idea.ed.gov

Kendall, P. C., Hedtke, K. A., & Aschenbrand, S. G. (2006). Anxiety disorders. In D. A. Wolfe & E. J. Mash (Eds.), *Behavioral and emotional disorders in adolescents* (pp. 259–299). New York: Guilford Press.

Morris, T. L., & March, J. S. (Eds.). (2004). *Anxiety disorders in children and adolescents* (2nd ed.). New York: Guilford Press.

Silverman, W. K., & Kurtines, W. M. (2001). Anxiety disorders. In J. N. Hughes, A. M. LaGreca, & J. C. Conoley (Eds.), *Handbook of psychological services for children and adolescents* (pp. 225–244). New York: Oxford University Press.

Telch, M. J., Smits, J. A., Brown, M., & Beckner, V. (2002). Treatment of anxiety disorders: Implications for medical cost offset. In N. Cummings, W. T. O'Donohue, K. E. Ferguson (Eds.), *The impact of medical cost offset on practice and research: Making it work for you* (pp. 167–200). Reno, NV: Context Press.

Vasey, M. W., & Dadds, M. R. (Eds.). (2001). *The developmental psychopathology of anxiety.* New York: Oxford University Press.

Vasey, M. W., & Ollendick, T. H. (2000). Anxiety. In A. J. Sameroff, M. Lewis, & S. M. Miller (Eds.), *Handbook of developmental psychopathology* (2nd ed., pp. 511–529). New York: Kluwer Academic/Plenum.

Wilmshurst, L. (2005). *Essentials of child psychopathology.* Hoboken, NJ: Wiley.

Applied Behavior Analysis

Applied behavior analysis is a methodology for systematically applying the principles of learning theory to develop interventions that will improve socially significant behaviors to a meaningful degree, as well as demonstrate that the interventions employed are responsible for the improvement in behavior. Applied behavior analysis has been repeatedly demonstrated to be a highly effective approach across a wide range of problems and environments, including education, mental health and mental retardation, parent training, environmental management, and organizational management. Applied behavior analysis is a specialty used by various professions. It is not regulated by most states, except as part of psychology or other established professions, although there have been a few attempts to recognize trained and qualified behavior analysts over the years. Most recently the nonprofit Behavior Analyst Certification Board has promoted a national certification program to identify and credential qualified practitioners and trainers.

Principles and Techniques

Several terms besides *applied behavior analysis* have been used to describe intervention methods based on behavioral learning theories, including *behavior modification, behavior therapy*, and others. Although sometimes used interchangeably, there are possible distinctions made. Applied behavior analysis is used

most often for the orientation that derives predominantly from Skinnerian operant conditioning and follows a radical behavioral philosophy. Although other behavioral orientations often utilize operant principles to differing degrees, they typically place a greater emphasis on classical conditioning processes (the neo-behavioristic mediational model) or cognitions and perceptions as targets for change (social learning theory and cognitive behavior modification) than does applied behavior analysis.

Operant conditioning eschews hypothetical mental constructs as explanatory contracts. Behaviors are viewed as being selected by environmental consequences, much like adaptive changes are in Darwinian evolution, rather than being emitted to serve some future purpose. That is, a child does not cry in order to attract attention but cries because crying has resulted in reinforcing consequences in similar situations in the past (unless, of course, the crying is in response to actual physical discomfort).

The central precept of applied behavior analysis is that behaviors are under the control of environmental stimuli. Functional relationships are described by a three-term contingency that consists of antecedents, responses (behaviors), and consequences. At times these are referred to as the ABCs of behavior. Early uses of behavior analysis focused primarily on changing behaviors through manipulating consequences. All consequences are seen as directly influencing whether behaviors will recur in the future. Consequences can be grouped into three main types: reinforcing, punishing, or neutral stimuli.

The first class of consequences, *reinforcers*, consists of events that increase the future probability of a behavior they immediately follow. These include events that strengthen behaviors when they are presented following the behavior, such as food, attention, or social praise. This operation is referred to as *positive reinforcement*. For example, a child may learn to apologize because the apology consistently is followed by parental praise. Behaviors also can be strengthened through the removal of an aversive (negative) stimulus following a behavior. This operation is termed *negative reinforcement*. This would be the case if a child learns to apologize if the apology terminates (or avoids) being scolded by his or her parents.

Reinforcers can either be biologically preestablished (*primary reinforcers*), such as food or water, or can acquire reinforcing properties through careful pairing with primary reinforcers (*conditioned reinforcers*).

Most reinforcers are differentially effective with different people rather than being universal. This is particularly true of conditioned reinforcers such as praise or tokens. The *schedule of reinforcement* used to deliver reinforcers also is important. Early in the process of strengthening a behavior, reinforcers are typically delivered on a continuous schedule, where a reinforcer is given each time the response occurs. Later in the process, various schedules are used that deliver reinforcers more intermittently and help increase resistance to extinction and/or produce specific types of responding patterns.

Punishment consists of two operations, but these weaken the likelihood of behaviors recurring. In *positive punishment*, the presentation of the consequence following the behavior results in weakening future occurrences of a behavior. An example might be a brief swat on the bottom when a young child chases a ball into the street. Removing a reinforcing stimulus contingent upon a behavior also can weaken it. This is *negative punishment* and includes, for example, taking away television-watching privileges for a short time to weaken a child's lying behavior. Applied behavior analysts typically advocate negative punishment as a more appropriate reductive method in most cases.

Behavior also may be decreased through the use of *extinction*, where the connection between a response and its maintaining consequences is discontinued, leading to a progressive decline in the rate of a previously reinforced response.

Antecedents, which are the front end of the three-term contingency, influence behaviors largely through a prior history of differential association with reinforcing or punishing consequences. When a child's behavior has been reinforced previously in a situation, the likelihood of that behavior is heightened under similar circumstances. It is lowered in situations in which reinforcement has been consistently withheld or punished. For example, a child whose father consistently buys him candy in a store when he whines is likely to exhibit similar behaviors in the future when shopping with his father. With his mother, who does not "give in," the child will learn not to whine.

Another type of antecedent stimulus used programmatically includes visual, verbal, or physical *prompts* given to increase the likelihood a child will respond appropriately to the given situation. For example, pictures of objects may be placed with letters to assist a child in learning letter sounds. These can then be

removed either abruptly or through a gradual process called *fading. Modeling* can be seen as a special form of prompting in which someone demonstrates a desired behavior to increase its likelihood.

When a behavior is not present in the individual's repertoire, the procedure of *shaping* or *successive approximations* may be used. Shaping involves reinforcing progressively closer approximations of the desired behavior. For example, in teaching new words to a young child, the child is reinforced for vocalizations that are increasingly more like the desired word. As the sequence progresses, the word must be more and more like the target word for the child to receive reinforcement.

Particularly important behavioral concepts for instructional applications include *discrimination*, *generalization*, and *concept formation*. Behaviorally, concept formation occurs when the same response occurs to a group of discriminably different objects that have some aspect in common (as well as responding differently to other classes of stimuli). Concept formation involves both generalization and discrimination: generalization within classes and discrimination between classes. Thus, a child who responds "dog" to different examples of dog but not cats or other animals is exhibiting concept formation.

Many other effective techniques have evolved over time from this relatively small set of basic principles. In addition to those just mentioned, these include procedures such as *overcorrection*, *token economies*, *time-out*, *response cost*, *self-monitoring*, and *task analysis*.

Methodological Practices

Although numerous step-by-step models have been proposed for developing, implementing, and evaluating applied behavior analytic interventions, each generally incorporates at least the following components:

1. Selecting a behavioral excess or deficit for change

2. Establishing a method for measuring behaviors

3. Measuring the baseline (current) level of performance

4. Specifying goals and objectives

5. Designing and implementing interventions to teach or strengthen behaviors/skills and/or to reduce excessive behaviors

6. Continuously measuring the behaviors to determine the effects of the intervention

7. Modifying the intervention based on ongoing measurements

In measuring behavior, applied behavior analysts focus on observable behaviors recorded using techniques such as direct observational recording (event, duration, or latency recording using continuous, time sampling or interval recording) and analysis of permanent products. In recent years, the practice of *functional behavioral assessment*—where the functions or maintaining factors of an existing behavior are determined experimentally prior to designing an intervention for implementation—has become increasingly more important.

A particularly important aspect of applied behavior analysis is the demonstration that the intervention used, rather than some unspecified extraneous variable, caused the behavior change. Applied behavior analysts accomplish this primarily through within-subject experimental designs that focus on the functional relationships between intervention changes and changes in target behavior. Typical experimental designs include reversal (ABAB) designs, multiple baseline designs, multiple-treatment designs, and changing criterion designs. The generality of effects is established through replication across subjects over time.

Historical and Current Applications

The field of operant conditioning, or the experimental analysis of behavior, began to grow exponentially during the 1930s. At that time B. F. Skinner built upon the basic concepts of radical behaviorism espoused by John Watson. Within a few years early attempts were made to apply Skinner's model of operant conditioning to human behavior and development. This was accompanied by a number of research programs, mostly in institutional settings. That work used arbitrary responses, as in animal operant research, to investigate the applicability of operant conditioning to humans. Concurrent with this, a few psychologists began working with the mentally retarded and mentally ill in clinical settings. The often-cited first human application, published in 1949, demonstrated that squirts of milk could be used to establish arm raising in a profoundly retarded 18-year-old man. Examples of other early applications included use of psychiatric nurses to change behaviors in psychotic patients and early attempts to change problem behaviors in preschool children.

During the late 1960s the field became increasingly focused on the application of operant principles to socially important problems rather than artificial laboratory research. The *Journal of Applied Behavior Analysis* was founded in 1968 as a primary outlet for this research. In the initial volume, Donald Baer, Montrose Wolf, and Todd Risley proposed seven defining dimensions of applied behavior analysis. When the same authors reviewed the status of these dimensions almost 20 years later, they concluded that the dimensions continued to be relevant. First, the behaviors or stimuli studied must be selected for applied (practical) rather than theoretical significance. They must need improvement and be measurable (behavioral). It is essential that the factors responsible for the occurrence or nonoccurrence of the behavior be established (analytic). The behavior change procedure used needs to be described in terms of the relevant principles (conceptual systems) and must be completely identified and described to permit replication of the interventions (technological). Behavioral techniques used must produce significant practical effects (effective). Lastly, those effects must be stable over time and situations, or they must extend to untrained responses (generality).

Since the 1950s applied behavior analysis has become an increasingly important applied field of psychology that has been used to modify behaviors across numerous areas of human functioning, including mental health and mental retardation, general and special education, organizational performance management, environmental management, and behavioral medicine.

Applied behavior analysis has been extended across many areas of education over the past few decades. Besides the obvious clinical applications for individuals with disabilities, a number of instructional methods used across regular education, special education, or adult and higher education have been grounded in applied behavior analytic principles. These include *direct instruction*, *precision teaching*, and *personalized instruction*. Although there have been numerous studies demonstrating the effectiveness of these approaches, they have not been widely adopted. In fact, numerous appeals have been made to have education be more accepting of applied behavior analysis as a basis for instructional innovations.

Ronald A. Madle

See also Behavior Modification; Cognitive Behavior Modification; Operant Conditioning; Reinforcement; Token Reinforcement Programs

Further Readings

Alberto, P. A., & Troutman, A. C. (2006). *Applied behavior analysis for teachers* (7th ed.). Upper Saddle River, NJ: Prentice Hall.

Austin, J., & Carr, J. E. (Eds.). (2000). *Handbook of applied behavior analysis.* Reno, NV: Context Press.

Baer, D. M., Wolf, M. M., & Risley, T. R. (1987). Some still-current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis*, *20*, 313–327.

Cooper, J. O., Heron, T. E., & Heward, W. L. (1987). *Applied behavior analysis.* Columbus, OH: Merrill.

Martens, B. K., Witt, J. C., Daly, E. J., III, & Vollmer, T. R. (1999). Behavior analysis: Theory and practice in educational settings. In C. R. Reynolds & T. B. Gutkin (Eds.), *The handbook of school psychology* (3rd ed., pp. 638–663). New York: Wiley.

Skinner, B. F. (1987). Whatever happened to psychology as the science of behavior? *American Psychologist*, *42*, 780–786.

APTITUDE

Aptitude can be defined as individual differences that are related to subsequent learning during a fixed time frame. The learning or acquisition of knowledge or skills can occur in a formal intervention (training or education) or in an informal setting (experience or mentoring). This definition can be fruitfully narrowed by specifying the domain of individual differences (cognitive, noncognitive) and the type of relationship (i.e., rate, accuracy) with learning. Numerous individual differences are related to learning, including cognitive abilities, personality traits, interests, and values. All of these can be considered aptitudes, and all can be evaluated in work, school, or avocational contexts. Aptitudes are discussed most commonly in reference to cognitive abilities within a formal educational or training context where the learning is labeled achievement.

However, this general definition defines aptitude mainly by its relationships with learning outcomes. If one also considers the nature of an aptitude, one finds two fundamentally different underlying definitions of aptitude. The failure to differentiate between these two conceptualizations can be the source of much confusion. The first definition proposes that aptitude comprises individual differences that develop from a combination of innate and environmental influences. Aptitude reflects a current repertoire of behaviors and behavioral tendencies that predict subsequent learning.

The second and quite different definition proposes that aptitude comprises individual differences that are innate and largely unchanging and immutable within normal circumstances. That is, aptitude is raw material that either facilitates learning or is actively used to learn. In the first definition one might say that a field had a strong aptitude for growing crops even though this is partially a function of having been covered with artificial fertilizer. The second definition would characterize the aptitude of the field based only on the fundamental and natural composition of the soil. The expected patterns of empirical findings that would be consistent with each definition differ.

The first definition frames aptitude as a reflection of current capabilities and tendencies. An aptitude is current skill and knowledge. In this case aptitude is based on a person's innate potential as it has been developed through learning opportunities and environmental stimulation or deprivation. Under this definition, aptitude is not a clear window into innate talent and will predict future learning but may also change itself due to new learning experiences. This means that aptitude scores may show mean increases, variance changes, and rank order changes after learning has occurred. For example, an algebra exam is likely to predict learning of trigonometry. Algebra scores at Time 1 will correlate with trigonometry scores at Time 2. Trigonometry scores are also likely to increase from Time 1 to Time 2. Finally, if some students are fundamentally more innately skilled at math but have not previously had the opportunity to receive good training or engage in deliberate practice, nonchance rank order changes, and a decrease in the variance on the measure of algebra knowledge and skill between Times 1 and 2 may also be observed. The argument that the algebra test is an aptitude for learning trigonometry would be strengthened if it predicts Time 2 trigonometry scores above and beyond what trigonometry knowledge at Time 1 predicts.

In the second definition one would expect that aptitude would remain unchanged by subsequent experiences, intervention, and learning. One researcher, Carroll, proposed that the ideal case for specifying an aptitude could be framed with six conditions. At Time 1 there would be (1) meaningful differences on the aptitude measure but (2) no valid variance on the achievement measure. Condition 2 also results in the implicit but unstated condition that achievement at Time 1 is uncorrelated with achievement at Time 2. (3) The aptitude measure would not be correlated with the achievement measure at Time 1 (prelearning event), and (4) the means and rank order of people on the aptitude measure would not change from Time 1 to Time 2. (5) However, the achievement measure would increase from Time 1 to Time 2, and (6) scores at Time 2 would become correlated with the aptitude measure. In other words, one would see an increase on the achievement measure from no knowledge to differences in knowledge among people, and the unchanging aptitude measure would predict these increases and differences. A few situations generally fit these conditions. Two clear examples are musical training and learning a foreign language, where students have not typically been exposed to the instrument or language and the predictors of learning are not strongly influenced by the educational training. However, in most cases, the achievement domain is based on, or connected to, an area of previous learning (e.g., calculus after having studied algebra or graduate-level psychology after having studied introductory psychology). In addition, aptitude is not unchanging and static. Aptitude, achievement, and ability are difficult to cleanly distinguish from one another.

Aptitude, Achievement, and Ability

Cognitive aptitude tests are commonly contrasted with achievement or ability tests. In nearly all cases, the distinction among the three is more semantic than substantive. Although cognitive tests can measure a very wide range of different abilities, they still measure acquired knowledge and skill; they all assess ability. A person's ability is the extent to which that person is able to correctly do the tasks directed toward a goal. Ability is simply a person's current capabilities. Therefore, aptitude is the potential to attain ability, and achievement is an improvement in ability. Current ability and aptitude levels are a function of prior achievements. Scores reflect gains in learning from some point in time. Finally, prior learning is consistently a good predictor of subsequent learning, and so the tests also evaluate aptitude. Therefore, the distinction among ability, achievement, and aptitude is often artificial.

A more productive taxonomy begins by considering four dimensions: the breadth of material sampled, curriculum represented, recency of learning sampled, and purpose of the assessment. A measure with a broad sampling would measure behaviors from a number of different domains. Tests with multiple different subtests, like the Wechsler Adult Intelligence Scale (WAIS), are

a good example of broad breadth. Curriculum represented is the extent to which a measure is linked to specific training or educational programs instead of more general or informal learning. The extent to which the curriculum represented in the measure reflects the experiences of the people being evaluated is sample dependent. Recency of learning refers to whether the measure captures more historical learning versus recently taught material. The recency of learning is dependent on the age and background of the sample being assessed. The purpose refers to the applied use of the measure, for example, prediction, assessment, training evaluation, placement, or counseling.

The combination of high breadth, a lack of formal curriculum representation, historical learning, and the purpose of evaluating current capacities would characterize an ability measure. In contrast, an assessment with low breadth, for example, a specific course or educational year's curriculum that was recently taught for the purpose of measuring student learning during the course, would typify an achievement measure. An aptitude measure might take the form of modest breadth and relatively recent learning in a specific curriculum to be paired with the purpose of assessing potential for the same domain. Alternatively, if one is uncertain about the specific nature of the learning domain (e.g., college major is not fixed) or wishes to assess potential for a broad array of domains (e.g., a liberal arts education), one might measure aptitude with high breadth, for example, a general or core curriculum from relatively recent learning experiences.

The taxonomy of four different characteristics of a cognitive measure illustrates the ever-present bandwidth fidelity trade-off. In a given period of time there is always a trade-off between the breadth and precision of measurement. Take, for example, an introductory psychology course. Using a 2-hour final exam, the instructor can either do a very broad but low fidelity assessment of a student's learning in the course or gain more precision about a narrower aspect of introductory psychology. The decision about what to evaluate and in what detail should be based on the goals of the assessment.

Correlations Between Aptitude and Criteria

Both cognitive and noncognitive aptitude measures have been shown to be predictive of learning and other accomplishments. Cognitive and noncognitive measures predict, to a moderate degree, subsequent acquisition of knowledge and skill in laboratory studies. These relationships hold for tasks that range from simple learning and memory to complex cognitive skills. In the educational domain, aptitude tests predict educational outcomes both early and later in school. A large literature has demonstrated the relationship between aptitude measures and performance in higher education. In the work domain, aptitude measures are associated with success in organizational training programs.

Nature and Nurture in Aptitudes

Research that has attempted to model the development of knowledge and skill suggests that personality traits, interests, and cognitive abilities are associated with the amount and domains of learning a person acquires. A person's ability and preferences appear to influence the direction of the educational and work careers as well as the types of things the person achieves. Both cognitive and noncognitive types of aptitude have a widespread influence on people's lives, and research indicates that biology plays a role. Research on aptitude measures has demonstrated that, as currently assessed, these individual differences have strong heritable and environmental components. Differences across people appear to be due to a combination of genetic and environmental factors. This also appears to be true for personality traits. People are predisposed to have certain behavioral tendencies, but the expression of these tendencies is also strongly influenced by the environment. New research has begun to identify brain structures that are associated with cognitive aptitudes. Specifically, the lateral prefrontal cortex appears to be implicated in general cognitive ability. However, this research takes us full circle to the two initial definitions of aptitude. Despite clear biological links with aptitude, this research does not demonstrate that differences among individuals, groups, or genders are necessarily due to an innate source.

Nathan R. Kuncel and David M. Klieger

See also Aptitude Tests; Individual Differences; Learning Style

Further Readings

Conard, M. A., (2004). Aptitude is not enough: How personality and behavior predict academic performance. *Journal of Research in Personality*, *40*(3), 339–346.

Hooper, S. R., Wakely, M. B., de Kruif, R., & Swartz, C. W. (2006). Aptitude–treatment interactions revisited: Effect of metacognitive intervention on subtypes of written expression in elementary school students. *Developmental Neuropsychology*, *29*, 217–241.

Nicholson, S. (2005). The benefits of aptitude testing for selecting medical students. *British Medical Journal*, *331*, 559–560.

APTITUDE TESTS

Aptitude can be defined as any characteristic that affects an individual's response to instruction or treatment. An *aptitude test* is intended to measure characteristics that influence the likely or potential response to environmental stimuli. Test outcomes are often used to predict future performance. Common characteristics associated with aptitude tests include intelligence, career skill/interest, and personality. Aptitude is typically identified as a characteristic of an individual. However, success is defined by both the characteristics of the individual and the ecology in which the individual exists. The aptitude for high academic achievement cannot be realized without the necessary experiences to promote learning and achievement.

Aptitude can be conceptualized as readiness to thrive. The outcomes from college aptitude tests, such as the Scholastic Aptitude Test, are used to guide admission decisions. They provide information about the student's capacity to thrive and learn academic material presented in higher education. The outcomes from career-related aptitude tests, such as the Strong Interest Inventory, provide information about individual characteristics that influence potential responses to job training.

The first aptitude tests were developed by Sir Francis Galton in the late 19th century to measure intelligence. The approach was later popularized by Alfred Binet and his colleagues when they developed and introduced the Binet Intelligence test. The Binet was used in France to identify those students who were most likely to benefit from public education. The underlying assumption associated with the use of the Binet was that there were individual differences that affect the potential to learn, and public financing should be spent to educate the students with the most potential. The assumption of individual differences endures today; however, intelligence tests have gone through much refinement through the years. In addition,

the success of such tests spawned various types of aptitude tests to address the diverse measurement needs of practitioners.

Aptitude tests vary in the number and specificity of the characteristics measured. Some measure broad aptitudes, such as general intelligence, scholastic ability, or personality; others target narrow aptitudes, such as mathematical or mechanical abilities. These tests are widely used in schools, counseling centers, and clinics as one piece of information within a comprehensive assessment. There are several categories of aptitude tests, and the following sections provide background information on development of aptitude tests and practical information on various types of aptitude tests, including intelligence tests, scholastic achievement tests, personality tests, and career tests. Finally, historical and present uses of aptitude tests are discussed

Intelligence Tests

As briefly mentioned in the previous section, testing of mental capacities began in the late 19th century by Sir Francis Galton. The foundation of his intelligence theory was the knowledge that people interacted with the environment through the five senses. This led Galton to theorize that the most intelligent people had innate abilities for enhanced sensory discrimination and retention of information. Galton opened a laboratory and collected a large amount of information at public gatherings, including the World's Fair. He made his measures available to the public; however, the tests proved to be problematic. The first practical intelligence test was developed and disseminated around 1905 by Alfred Binet and his colleagues in France. The Binet–Simon Scale was commissioned by the French government to screen children for mental retardation. It was the first test to consider cognitive development in children, rank items by level of difficulty, and use a standard form of instructions throughout the test. The test was brought to the United States in 1908 by Henry H. Goddard and then adapted in 1916 by Lewis Terman at Stanford University. It is in this revision that the test was renamed the Stanford–Binet, and it remains one of the most popular intelligence tests given in schools.

After the introduction of intelligence tests to the United States, the Army created the first group-administered intelligence test. This test was dubbed the Army Alpha and comprised a series of questions

with true-or-false responses that covered content which included vocabulary, arithmetic, sentence structure, general knowledge, and practical judgment. The test was given to incoming soldiers during the final years of World War I in order to place soldiers in officer or infantry tracks. The Army also developed a nonverbal version of the test, the Army Beta, to provide the testing opportunity for those who did not speak or write in English and minimize misclassification of the non-English-speaking soldiers as less intelligent and therefore only appropriate for infantry service. This large-scale testing endeavor resulted in other large-scale testing programs and evolved into the Army's present-day testing program, the Armed Services Vocational Aptitude Battery.

Many other intelligence tests followed the Stanford–Binet and Army tests. One of the most popular is the Wechsler intelligence tests, which is a series of four intelligence tests for use with children and adults. The series includes a preschool version, child version, adult version, and a brief version for use with individuals across the testing age span. David Wechsler developed the first version, the Wechsler–Bellevue Intelligence Scale, in 1939 by combining subtests from various other intelligence tests, including the Stanford–Binet, Army Alpha, Army Beta, and Kohs Block Design Test. The most recent version geared toward children, the Wechsler Intelligence Scale for Children–IV (WISC–IV) was released in 2003, and the Wechsler Adult Intelligence Scale (WAIS, 1997) is in its third revision.

A more recent addition to the intelligence testing community is the Woodcock-Johnson–III Tests of Cognitive Abilities (WJ–III Cog, 2001). The WJ–III Cog development is theoretically based on the Cattell-Horn-Carroll theory of intelligence (CHC theory), which is the most widely accepted theory of intelligence within the field of psychology. The original Woodcock-Johnson Psycho-Educational Battery was published in 1977 and was the first co-normed battery of cognitive abilities, scholastic achievement, and interests for individuals of a wide age range. The first revision in 1989 separated the larger test into two distinct batteries, Tests of Cognitive Abilities and Tests of Achievement. The two test batteries share a large normative sample, which provides greater validity when making comparisons between the two tests within individuals, such as analysis of intelligence and achievement discrepancies for special education eligibility decisions.

Academic Achievement Tests

Intelligence tests are often accompanied by tests of academic achievement that provide a measurement of specific academic skills in content areas such as reading and math. The purpose of achievement tests is to measure individual academic strengths and difficulties. Outcomes can be used to describe and predict performance within the academic context. There are comprehensive achievement tests that cover a variety of academic skills (i.e., reading, math, social studies, science), and there are achievement tests that aim to measure a single academic skill (narrow-band tests) in more detail than is possible with a comprehensive test.

Achievement tests are administered in many contexts that include both group and individual administration. Group-administered tests include state accountability tests, such as the Iowa Tests of Basic Skills and college entrance exams (e.g., Scholastic Aptitude Test [SAT], Graduate Record Exam [GRE], and American College Testing [ACT] program). Individually administered tests are generally used within the context of special education eligibility assessments and include the Woodcock-Johnson–III Test of Achievement, Stanford Achievement Test, Wechsler Individual Achievement Test, and many narrow-band assessments (e.g., Gray Oral Reading Test–IV).

Individually administered achievement tests and state accountability tests are not often considered aptitude tests because the purpose of the test is quite different from the defined application of aptitude tests. Individual achievement tests are used to describe student strengths and difficulties at the time the test is given. These data are then used to make decisions about the provision of supplemental academic services, such as classroom interventions or special education, but the data are not used to predict future scholastic attainment. The data from state accountability testing are also used for the purpose of description of current academic achievement but are not often used to predict student achievement in the future.

College aptitude tests represent one of the largest and best-known applications. They are used to guide admissions to colleges and universities at both undergraduate and graduate levels. Students begin to prepare for college entrance exams as early as middle school, but more frequently the preparation begins in the sophomore year of high school. The Preliminary Scholastic Aptitude Test (PSAT) is a preparatory test, similar in format to the SAT, which provides students

with the opportunity to practice taking a college entrance exam without the pressure of sending scores to colleges and universities and also acts as an assessment for the National Merit Scholarship competition. Scores on the PSAT can be used to predict future performance on the SAT and as the first step in the evaluation of students for the National Merit Scholarship.

The PSAT is followed the next year with the SAT and/or the ACT. Colleges and universities across the country use these academic achievement scores as part of the admission criteria. The SAT was developed by the Educational Testing Service of Princeton at the request of the College Board and is one test in a series of large-scale assessments that also includes the PSAT, the high school Advanced Placement tests, and the Graduate Record Examination (GRE). The SAT general test consists of three subtests that measure critical reading, mathematics, and writing. Subject tests are also available and fall within five major subject areas: English, history, science, mathematics, and languages. Colleges do not often require subject tests but encourage students to take one or more of the subject tests for proper placement within tracked courses, such as mathematics.

The ACT is the largest competitor of the SAT. It was developed in 1959 by E. F. Lindquist with the University of Iowa Testing Programs. The test was developed to assess the academic skills needed to succeed at the college level and measures content that is regularly taught in high school curricula. The test comprises five subtests that assess reasoning and problem solving in English, mathematics, reading, and science. The writing portion of the test is optional. Similar to the SAT, the ACT provides academic achievement scores, but it differs from the SAT with the addition of career interest items given in combination with the academic content. The final score report provides information about academic and career interests of the individual for college and career planning purposes.

Similar to the undergraduate college admission process, graduate programs also require aptitude testing. Graduate programs require the GRE General Test and often recommend GRE Subject Tests as a measure of achievement in a specific area of undergraduate study. The GRE General Test measures skills in verbal and mathematics reasoning, critical thinking, and analytical writing that are acquired throughout the many years of education and are not specific to any field of study. The GRE Subject Tests are available for eight specialty topics: biochemistry, cell and molecular biology,

biology, chemistry, computer science, literature in English, mathematics, physics, and psychology. Professional schools (i.e., medical school and law school) do not require the GRE but instead require tests that measure aptitudes required by the profession. The Law School Admission Test (LSAT) is administered by the Law School Admission Council and is a test of logical and analytical reasoning, reading comprehension, and analytical writing skills. Scores on the LSAT are used as a criterion in law school admissions decisions because they positively correlate with first-year law school grades. The Association of American Medical Colleges developed the Medical College Admissions Test (MCAT). The MCAT measures verbal and writing skills and knowledge in biological and physical sciences and is an admissions requirement at almost all medical schools in the United States.

Career Tests

Aptitude tests are not limited to scholastic abilities and skills. Whereas some individuals may understand their personal skills and interests well enough to choose a career easily, others many have trouble with this process and need guidance. Career aptitude tests were developed to help guide individuals in the process of career decision making.

The Differential Aptitude Test and General Aptitude Test Battery were both designed to assess various cognitive and physical abilities and provide guidance to individuals making vocational decisions. They are designed to reflect differences among the various aptitudes, which include verbal and numerical reasoning, mechanical and clerical aptitudes, manual dexterity, and motor coordination. The interpretation of scores on these tests is directly related to various occupations. Individuals are provided with lists of career and vocational environments that best match the pattern of aptitudes reflected in their test scores.

Another type of career-related aptitude test is the interest inventory. The Strong Interest Inventory was developed in the 1920s by E. K. Strong and is based on John L. Holland's vocational choice theory. Holland theorized that people and work environments could be categorized into six basic types: Realistic, Artistic, Social, Investigative, Enterprising, and Conventional. Holland believed that people are most happy in their work when personality, skills, and interests match the work environment. Very few individuals and work environments can be categorized as a pure

type; therefore, Holland theorized that each could be categorized as a combination of three of the six basic types. With Holland types as a foundation, the Strong Interest Inventory was developed through discussions with individuals who were happy with their work environments, and the 219 items subsequently created ask clients to rate their interests in careers, skills, and job-related tasks. The reported scores on this test are quite different from those received from intelligence and achievement tests because there are no incorrect answers on this test. The client receives a three-letter type code that corresponds to Holland's six basic types and is related to the careers of those who were interviewed in the development of the test. The report also provides a list of careers the individual might enjoy and, in the case of high school or college students, a list of college majors the individual might consider.

While the Strong Interest Inventory is quite popular, the Myers-Briggs Type Indicator (MBTI) is one of the most popular of all career-related aptitude tests. It is often used in the context of career guidance, but it was developed as a test of personality types. The MBTI was developed by Katharine Cook Briggs and Isabel Briggs Myers in the mid-20th century and is based on the personality theories of Carl Jung. The MBTI is composed of 93 forced-choice items whose responses are then sorted into four types. The four bipolar scale types, known as dichotomies, are introversion/extroversion, sensing/intuition, thinking/feeling, and judging/perceiving. The test results in a four-letter code that refers to the preferences of the individual on each of the four dichotomies. For example, a person might receive a code of ESTJ, which places that person on the extroverted side of the continuum, sensing is considered stronger than intuiting, thinking is stronger than feeling, and the individual is more of a judger than a perceiver.

The Strong Interest Inventory and MBTI are often used in conjunction as a battery for career guidance assessment. Whereas the Strong Interest Inventory has a strong support for reliability and validity in the research literature, the MBTI has much less support and should not be used for decision-making purposes but rather for personality and career exploration in conjunction with other career aptitude measures.

Personality Tests

Large-scale intelligence testing began during World War I with the need for U.S. Army personnel to

properly place new soldiers into positions that best suited their interests and skills. The need for placement of World War I soldiers also influenced the development of personality tests. Robert Woodworth developed a series of forced-choice items that were used to screen soldiers for potential mental health problems such as excessive anxiety. Others followed with their contributions to personality assessment, but personality inventories were not widely used in the general population until Robert G. Bemeuter developed his Personality Inventory in 1930. Personality assessment became even more widespread with the introduction of the Minnesota Multiphasic Personality Inventory (MMPI), a personality inventory that measures emotional and social disturbance. The MMPI was developed by Starke Hathaway and J. C. McKinley in 1943 to aid in diagnosis of psychiatric patients and is the mostly widely researched instrument in print. The MMPI was developed empirically by asking patients diagnosed with a wide range of psychiatric disorders and a large number of control participants to respond to a series of true/false statements. They identified 504 items that differentiated between the 221 clinical patients and the pool of 1,508 control participants. The MMPI was revised in 1989 with updated items and a more representative normative sample. The MMPI–2 contains 567 items, 10 clinical scales, 3 validity scales, several content and supplementary scales, and a list of critical items all developed by various researchers. The 10 clinical scales are Hypochondriasis (Hs, 1), Depression (D, 2), Conversion Hysteria (Hy, 3), Psychopathic Deviant (Pd, 4), Masculine-Feminine Interest (MF, 5), Paranoia (Pa, 6), Psychasthenia (Pt, 7), Schizophrenia (Sz, 8), Hypomania (Ma, 9), and Social Introversion (Si, 0). The clinical scales retained the original names even though many of these terms are not used in present-day psychology. Today the scales are commonly referred to by number. Recently, the Restructured Clinical scales were developed by Auke Tellgen to aid interpretation and provide greater alignment with the present categorical system used in psychology diagnosis. The MMPI–2 can only be used with adults 17 years of age or older. There is an adolescent version, the MMPI–A, but a version for childhood assessment is not available.

In contrast to the MMPI–2, which aims to measure disordered personality tendencies, other personality tests aim to measure normal personality traits. The NEO-Personality Inventory (NEO-PI) was developed

to measure the Five-Factor Model of personality. The five factors are Neuroticism, Extroversion, Openness, Conscientiousness, and Agreeableness. The Five-Factor model of personality is one of the most widely accepted trait theories of personality. Whereas the MMPI was empirically derived, the NEO-PI is based on theory, and items were chosen through the employ of elaborate statistical procedures. The NEO-PI is often used in employment settings, but there is some research to support its use in clinical diagnosis. Similar to the MMPI, the NEO-PI was standardized for use with adults only.

Children's personality inventories are also available and include the Child Behavior Checklist, the Children's Depression Inventory, and the Behavioral Assessment System for Children (BASC). However, these inventories are not referred to as personality tests because it is widely understood personality is not stable until early adulthood; therefore, temperament and behavior are measured prior to adolescence. Childhood assessments commonly include items for parents, the child's teacher, and self-report from the child. This multisource approach provides greater reliability and validity of the assessment. The BASC–2 (revised in 2006), developed by Cecil R. Reynolds and Randy Kamphaus, is a widely used assessment that combines parent, teacher, and student self-reports along with a direct behavioral observation system. This multisource, multimethod comprehensive assessment is used frequently in educational and clinical settings for the purpose of behavioral and mental health assessment of children.

There are a large number of personality inventories available for use with clinical, educational, and normal populations, but a description of each is outside the scope of this entry.

Application of Aptitude Tests

Whereas this entry takes a broader view of aptitude as put forth by Robert Snow, aptitude is often considered synonymous with intelligence. Throughout the mid-20th century, when psychological and educational research began to flourish, there was an increased interest in using intelligence tests to pinpoint cognitive difficulties and subsequently provide interventions to individuals based on the intelligence scores. Interventions targeted specific difficulties as indicated by the intelligence test. This process is referred to as an *aptitude by treatment interaction*

(ATI) approach. ATI was often used as a mechanism to diagnose and treat learning disabilities, but decades of research have concluded that ATI did not function to increase student cognitive skills, especially the academic skills valued in education, such as reading and mathematics.

Currently, the accepted use of aptitude tests is as part of a multisource, multimethod assessment where the aptitude test provides only one piece of the information used in the determination of intervention for an academic or behavioral difficulty. Aptitude tests function to predict future performance, and the addition of information from various sources (i.e., parent, teacher) and different assessment methods (i.e., behavioral observations, record review) increases the accuracy of future performance.

Implications

Aptitude is any characteristic that affects an individual's response to intervention. Although *aptitude* is often considered as synonymous with *intelligence*, the definition presented here provides a much broader view of aptitude that can include characteristics such as academic skills, personality traits, and career interests. Therefore, aptitude tests measure individual characteristics that influence responses to environmental stimuli and function to predict future performance.

The formal development of aptitude tests began in the mid-1800s and has since become a prolific psychological science that allows clinicians to predict future performance with increased precision. Historically, the use of aptitude tests was limited to an ATI framework, but the research has supported a broader use of these tests within a multimethod assessment. Aptitude tests can provide clinicians with information about individual characteristics that support readiness for treatment, intervention, or instruction. However, individual aptitudes do not necessarily translate into life success or the establishment of skills. They must be molded by complementary experiences. Those experiences might be formal, such as schooling, job training, or counseling. Those experiences might also be informal, such as interactions with family, friends, and daily problem solving.

In this age of prolific psychological research, tests are updated and new tests are developed. It is most important, when choosing an aptitude test, to understand the reliability and validity evidence, the intended

purpose of the test, and the population for which the test was developed.

Theodore James Christ and Nicole Skaar

See also Aptitude; Assessment; Measurement; Reliability; Validity

Further Readings

Ackerman, P. L., & Wolman, S. D. (2007). Determinants and validity of self-estimates of abilities and self-concept measures. *Journal of Experimental Psychology: Applied*, *13*, 57–78.

Reid, C. A., Kolakowsky-Hayner, S. A., Lewis, A. N., & Armstrong, A. J. (2007). Modern psychometric methodology: Applications of item response theory. *Rehabilitation Counseling Bulletin*, *50*, 177–188.

Stoeber, J., & Kersting, M. (2007). Perfectionism and aptitude test performance: Testees who strive for perfection achieve better test results. *Personality and Individual Differences*, *42*, 1093–1103.

ASIAN AMERICANS

The term *Asian Americans* refers to Americans, and immigrants living in America, whose ancestral heritage is linked to several Asian regions of the world: East Asia, Southeast Asia, and South Asia. The region of East Asia comprises a variety of countries like China, Japan, South Korea, North Korea, Taiwan, Hong Kong, Macau, and Mongolia; Southeast Asia consists of Cambodia, Indonesia, Laos, Malaysia the Philippines, Singapore, Thailand, and Vietnam; and the region of South Asia includes Bangladesh, India, Myanmar, Nepal, and Pakistan. Although Pacific Islanders (Hawaiians, Guamanians, and Samoans) may also be considered as Asian Americans, typically they are listed separately; therefore, they are not discussed here even though many of the issues discussed in this entry are applicable to the Pacific Islanders.

The Asian American population is not a homogeneous group. Yet, the term *Asian American* is used to categorize these diverse ethnic groups together. Although Asian Americans may share certain commonalities, they are an extremely diverse group of people from different countries and with rich cultures that are vastly different from each other. A true understanding of Asian Americans begins with their country's history, sociopolitical history in America, cultural practices and beliefs, language, conditions surrounding immigration, experiences after arrival in the United States, socioeconomic background, family situation, ethnic identity and assimilation, and length of residence in the United States.

Educational psychology has been interested in Asian Americans because they are a leading ethnic group with a long immigration history in the United States. Research with Asian Americans has focused on issues such as educational achievement, how they learn and develop in similar and dissimilar ways compared with White European Americans, and their psychosocial adjustment to mainstream American culture and schools. In the following paragraphs, Asian Americans are discussed along with their immigration history, Asian American subgroups, and Asian American issues.

Immigration History

The first Asian immigrants arrived in the United States more than 150 years ago; however, most Asian Americans arrived in the United States recently, after the adoption of the Immigration and Nationality Act of 1965. Prior to 1965, immigration from Asia had been barred by the National Origins Quota Act of 1924. In general, the Asian immigrants who moved to the United States as beneficiaries of the Immigration and Nationality Act of 1965 are characterized as highly professional and educated individuals. They arrived in the United States as skilled professionals, entrepreneurs, students, and family members of already naturalized Asian Americans. They migrated to the United States to seek better economic, social, and educational conditions. In contrast, Asian refugees, primarily from Southeast Asia, migrated to the United States as a result of war and political persecution and, thus, tended to be less educated and more economically disadvantaged.

The number of Asian Americans has grown more than eightfold since 1970. The U.S. Census Bureau estimated that there were approximately 12 million Asian Americans in the United States in 2005. The large population growth was the result of increased immigration from China, India, Korea, the Philippines, and other Asian areas after 1965 and the entry of more than 1.5 million Southeast Asian refugees following the end of the Vietnam War. The Asian American population is estimated to grow to 20 million by the year 2020.

Because of historical immigration patterns, Asian Americans tend to be highly concentrated in a few geographic regions. Currently, more than half of the Asian Americans reside in Hawai'i, California, and other West Coast states. In addition, the Asian American population is heavily concentrated in the major metropolitan areas such as Los Angeles, New York, San Francisco, Washington, D.C., and Chicago, where many ethnic enclaves exist. However, the patterns of these geographic concentrations differ greatly from one Asian ethnic group to another. Furthermore, more recently, the Asian American population has started dispersing gradually throughout different regions in the United States.

Asian Americans represent a number of diverse groups who differ in national origin, ethnic identity, language, culture, religion, socioeconomic background, experiences after arrival in the United States, and length of residence in the United States. For a better understanding of Asian Americans, it is essential to recognize the diversity of the Asian Americans today. Despite their diverse backgrounds, Asian Americans can be categorized into three main subgroups based on their geographic origins: East Asian, Southeast Asian, and South Asian. In the following section, the main characteristics of the three groups are explained.

Asian American Subgroups

East Asian

East Asia generally includes China, Japan, South Korea, North Korea, Taiwan, Hong Kong, Macau, and Mongolia. Among them, Chinese, Japanese, and Korean are the major East Asian ethnic groups found in the United States. Research on the other East Asian ethnic groups in the United Sates is very limited. Chinese, Japanese, and Koreans share the influence of Taoism, Confucianism, and Buddhism, which emphasize interpersonal harmony, social order, a holistic vision of life, and the importance of education. The traditional East Asian families maintain a hierarchical family structure and gender-specific roles. These three groups maintain a high level of educational attainment in the United States.

Despite commonalities among East Asian ethnic groups, they are quite different. Chinese Americans are the largest Asian ethnic group in the United States. Currently, there are more than 2.6 million Chinese Americans living in the United States.

Although Chinese immigrants were among the first Asians to settle in the United States in 1840s, their recent population growth is mainly due to the passage of the Immigration and Nationality Act of 1965. Accordingly, 71% of Chinese Americans in the United States were foreign born. In addition, 85% of the Chinese Americans speak a language other than English at home. Unlike the early Chinese immigrants who were peasants, the majority of recent Chinese immigrants are highly educated and affluent. This demographic shift in the Chinese American population after 1965 has resulted in the polarization of the Chinese community into the working and professional classes. The professional class tends to acculturate well to American society, whereas the working class tends to stay in ethnic enclaves and retain their Chinese traditions.

Similar to Chinese Americans, Japanese Americans were early Asian immigrants in the United States. Although Japanese Americans were the most populous Asian ethnic group between 1910 and 1960, its recent population growth is much slower than the population growth of other Asian ethnic groups in the United States. Accordingly, fewer Japanese Americans are foreign born (40%), and many Japanese Americans are third, fourth, and fifth generations. In addition, Japanese Americans are the oldest Asian American group, with a median age of 42.6 years. The majority of Japanese Americans speak English well. In fact, 53% of Japanese Americans speak only English at home. The poverty rate for the Japanese Americans was the lowest among all the Asian ethnic groups in the United States in 2000. Other data on educational attainment, occupation, and housing suggest that Japanese Americans have a higher level of acculturation compared with other Asian ethnic groups in the United States.

Compared with Chinese Americans and Japanese Americans, Korean Americans have a shorter immigration history in the United States. The majority of Korean immigrants arrived in the United States to seek better employment and educational opportunities after 1965. Many of these recent Korean American immigrants were well educated and from the middle class. The 2000 U.S. Census data indicate that 70% of Korean Americans were foreign born and the majority of Korean Americans speak Korean at home. Several research studies suggest that Korean Americans are more likely to maintain their native culture and participate in ethnic social networks. Their relatively short length of residence in the United States, and the fact that Korean Americans are a homogeneous ethnic

group (speaking one language) who retain their traditional value of education, may largely contribute to their relatively high levels of achievement. In addition, the majority of Korean Americans have strong affiliations with Korean churches, which are primarily Protestant (e.g., Presbyterian, Methodist, Baptist) or Roman Catholic. The Korean churches play a significant role in maintaining their ethnic social networks by providing support and resources for Korean Americans to endure the hardships of acculturation.

Southeast Asian

Southeast Asia generally includes the countries of Cambodia, Indonesia, Laos, Malaysia, the Philippines, Singapore, Thailand, and Vietnam. The most significant characteristic of Southeast Asians is that they are much more ethnically and culturally diverse than East and South Asians. In fact, Southeast Asians with the same national origins can differ greatly in terms of ethnicity, language, and religion. To better understand the Southeast Asian population, it is helpful to distinguish them based on their migration status: Southeast Asian *immigrants* and Southeast Asian *refugees*. More specifically, the purpose of their migrations and the national origins of these two groups have distinctive patterns.

The Southeast Asian immigrants are primarily from Indonesia, Malaysia, the Philippines, and Thailand. Except for Filipinos, who were the only Asians not barred from immigration by the National Origins Quota Act of 1924, most of the Southeast Asian immigrants started migrating to the United States after the enactment of the Immigration and Nationality Act of 1965. The majority of these Southeast Asian immigrants were highly educated and skilled professionals who migrated to the United States to seek better employment.

Among all the Southeast Asian immigrant groups, the Filipinos are currently the largest Southeast Asian immigrant group and the second largest Asian ethnic group in the United States. Many research studies have been done on the Filipino Americans, whereas little is known about the other Southeast Asian immigrant groups. According to these studies, the Filipinos are the most Westernized Asian Americans because of their colonization by Spain and the United States and because of the strong influence of Roman Catholicism. Although the majority of the Filipinos in the United States are foreign born, the Filipinos generally do not have serious language barriers. The Filipino family

structure is slightly different from those of other Asian groups because of the indigenous Filipino culture prior to the Spanish cultural colonization. The Filipino family is more egalitarian, where husband and wife share almost equal power in financial and family decisions. In addition, kinship relationships are highly valued in the Filipino family and play a significant role to provide support for many Filipino Americans in their early stage of adjustment to American society.

In contrast to the Southeast Asian immigrants, the majority of the Southeast Asian refugees arrived in the United States after 1975 following the withdrawal of the U.S. troops from Southeast Asia. They are the newest Asian ethnic group in the United States and their experiences of migration and settlement differ fundamentally from those of other Asian ethnic groups. The Southeast Asian refugees primarily include Cambodians, Laotians, Hmong, and Vietnamese. They migrated to the United States as part of the refugee resettlement program. As political refugees, many of the Southeast Asian refugees experienced traumatic departures from their native countries and spent years in refugee camps in the Philippines, Malaysia, and Thailand before they were admitted to the United States.

Except for the first wave of the Southeast Asian refugees who arrived in the United States in 1975, the majority of the Southeast Asian refugees were less educated and less exposed to Western cultures prior to their migration. In addition, the Southeast Asian refugees were among the youngest Asian ethnic groups in the United States, which is partially due to high birth rates with median ages of less than 20 years for all groups except the Vietnamese. The median age of Hmong in 2000 was 16.3 years. Because of their short history of residence in the United States, the majority of the Southeast Asian refugees do not speak English well. The Southeast Asian refugees are more likely than any other Asian ethnic group to speak a non-English language at home. Although the children in the refugee camps were taught the English language and Western culture, many of them lost opportunities to have a formal education. The 2000 U.S. Census indicates that educational attainment rates for Cambodians, Laotians, Hmong, and Vietnamese were much lower than those of other Asian ethnic groups.

The Vietnamese Americans are the largest Southeast Asian refugee group and the second largest Southeast Asian ethnic group after the Filipinos. Many Vietnamese Americans are ethnically Chinese

who had lived in Vietnam for generations. Therefore, their culture is strongly influenced by Buddhism, Confucianism, and Taoism. Buddhism also strongly influences the Cambodians and Laotians. The Cambodians and Laotians value harmony in interpersonal relations and keep close family relationships. Their societies are characterized by several layers of social classes. The Hmong, one of the ethnic groups from Laos, are strongly influenced by shamanism. Their strong belief in shamanism and their distrust of Western medicine have often created conflicts with social service systems in the United States. Hmong are also known for their strict patriarchal clan system. Many Hmong families maintain close family ties in the United States and geographical concentrations in states such as California and Minnesota. Their acculturation level is much lower compared with other Asian ethnic groups in the United States.

South Asian

South Asia generally includes Bangladesh, India, Myanmar, Nepal, and Pakistan. Although the first wave of South Asian immigrants arrived in the United States in the early 20th century, the majority of the South Asians migrated to the United States after the immigration law was reformed in 1965. Currently, over 1 million South Asians reside in the United States. The South Asian population in the United States consists of a large Asian Indian population, with smaller numbers of Pakistani, Bangladeshis, and other South Asian ethnic groups. The Asian Indians were the fourth largest Asian ethnic group in the United States in 2000. India, Pakistan, and Bangladesh, which were one big colony under British rule, share relatively similar ethnic and cultural backgrounds with the exception of religions. The major religions of the South Asians include Hinduism, Islam, Sikhism, Christianity, and Jainism. Whereas the majority of the Asian Indians are Hindus, the majority of the Pakistanis and Bangladeshis are Muslims.

Similar to other Asian immigrants, the South Asian immigrants consist of highly educated professionals. Many of the recent South Asian immigrants first arrived in the United States as students and later adjusted their status to permanent residents or naturalized citizens. The U.S. Census 2000 indicates that the three major South Asian ethnic groups had exceptionally high college-completion rates. A large number of South Asian immigrants are represented in professional occupations like medicine or other science-related fields. Because

of British colonization, the majority of South Asians speak English fluently. South Asian immigrants, especially Asian Indians, seem to have less trouble with acculturation to U.S. society compared with other Asian American ethnic groups. In fact, Asian Indians are more likely than other Asian immigrants to become naturalized citizens. However, despite their fluency in English and educational achievements, many South Asians suffer from personal and institutionalized racism in the United States. In fact, an increasing number of hate crimes against the South Asians have been reported since the aftermath of 9/11 in 2001. Because of the physical resemblance to people of Middle Eastern descent and the tradition of wearing a turban, Sikh Indians are unfortunately more likely to be targets of hate crimes.

Issues Facing Asian Americans

Model Minority Myth

The image of Asian Americans as a model minority started appearing in the American mainstream media in the 1960s. The term *model minority* was first used in the article "Success Story: Japanese American Style," written by sociologist William Peterson and published in the *New York Times Magazine* in 1960. A similar article using Chinese Americans also appeared in other news magazines in the same year. These articles highly praised the Japanese and Chinese Americans for overcoming hardships and discrimination and achieving success in American society. Thus, the model minority stereotype changed previously held negative images of Asian Americans. Asian Americans are still frequently viewed as a highly successful minority group in the United States. Asian Americans are portrayed as hardworking, highly educated individuals with few psychological problems; this portrayal leads to an inherent assumption that they are less of a burden on American society.

Data Supporting Stereotype

In support of the model minority view, Asian Americans share a great respect for the importance of education. For example, the 2000 census data showed that Asian American parents are more likely to have higher educational expectations for their children than are parents in other race groups. The academic successes of Asian American students have been documented numerous times. In 2000, the college enrollment rate

was 56%, which is higher than other racial groups, including Whites. The proportion of Asian students in Ivy League schools and other top universities in the nation also indicates their academic success in higher education.

In primary and secondary education, this trend also seems to persist. The proportion of Asian American students in gifted classes is much higher compared with their non-White counterparts. The proportion of children who had repeated a grade was the lowest for Asian Americans (2.6% for children 6 to 11 years old and 6.4% for children 12 to 17 years old). Fewer Asian American students between ages 12 and 17 were suspended from schools (5.3%) than White students of the same age span (9.0%). Furthermore, the high school dropout rate of the Asian American students in 2000 (4.0%) was the lowest among all races in the United States.

Data Refuting Stereotype

In sharp contrast to the model minority myth is the fact that the educational attainment data vary greatly by Asian ethnic groups, as discussed previously. One factor influencing these great discrepancies in educational attainment among Asian Americans is their family's migration status to the United States. Whereas many East Asians arrived in the United States as immigrants, most Southeast Asian immigrants arrived as refugees after 1975. Therefore, these two groups differ in terms of parents' level of education and economic status prior to immigration. Many East Asian children have parents with at least bachelor's degrees, whereas many Southeast Asian children have parents who have 9 years of education or less. Accordingly, the household income for East Asians is higher than average, and the poverty rate among Southeast Asian immigrants is higher than average. Educators need to recognize the group differences in educational attainment among Asian American students. Asian American children who have parents with little formal education and moved to the United States as refugees might need additional support at school and at home in order to progress successfully in school and successfully transition to adulthood.

Negative Effects of the Model Minority Myth

The model minority stereotype is a myth that is hurtful to Asian Americans and negatively affects society. Its oversimplification of the Asian American population is misleading and overlooks the great diversity among the Asian American ethnic groups, which is related to the stereotype being created in the 1960s before the recent influx of Asian immigrants and refugees. The increasing heterogeneity and complexity among Asian Americans suggests that the model minority stereotype is creating a great misconception about Asian Americans today.

Still, many non–Asian Americans may think that the model minority stereotype is positive and therefore beneficial to Asian Americans. For example, being considered smart and hardworking is good and may increase one's self-esteem. However, the pressure to succeed academically and to conform to the model minority stereotype has put an excessive amount of pressure on Asian American children and adolescents. Children and adolescents with such pressure from school and society may be more prone to difficulties with anxiety. In schools, Asian American students are often stereotyped as "model" or "successful" students who are hardworking and polite. As a result, the needs of the Asian American students may be frequently overlooked by educators, creating a series of events that may lead to academic and school disengagement and dropout. This becomes especially serious with the Asian American children who are foreign born or who are children of the recent refugees because they need extra support from teachers and other school professionals to adjust themselves to their new school environment. Teachers who believe the model stereotype might expect more from the Asian American students academically. Furthermore, because of the assumption that Asian American students are quiet and passive, teachers might treat them more strictly than they would treat other minority children when Asian American students behave contrary to their expectations.

Finally, in addition to negative education and schooling effects, the model minority myth is frequently associated with false assumptions about racism, poverty, and mental health. The related assumptions are that because of Asian Americans' educational success, their "privileged status" protects them from racism, poverty, and mental health difficulties. On the contrary, these beliefs are false. For example, Asian Americans have historically been the targets of extreme racist acts (e.g., Chinese Exclusion Act of 1882) and continue to experience racism today. One current form of racism is called "glass ceiling,"

which refers to a situation where people are prevented from advancing to higher positions in their careers. The fact that well-qualified Asian Americans are underrepresented in executive and managerial positions suggests the existence of the glass ceiling against the Asian Americans in workplaces. In schools, despite the diversity of Asian American children and adolescents, many Asian American children often experience racism in the form of fighting, name-calling, and teasing because of their race, English proficiency, accent, and foreign attire.

Racism and discrimination are generally exacerbated during economic and national crises. For example, the hate crime statistics reported a sharp increase in hate crimes based on religion and ethnicity/national origin in 2001 after the September 11 terrorist attacks. Asian American Muslims, such as Bangladeshis, Pakistanis, and Indonesians, and South Asians who were perceived to be Arabs or Muslims became the targets of these hate crimes. The hate crimes against Asian Americans included murder, physical assaults, vandalism of places of worship, death threats, and public harassment.

Finally, certain Asian American groups live in poverty and experience mental health difficulties, especially as a result of traumatic immigration experiences.

Identity Development

Identity development for the Asian American is more complex than for someone from the mainstream majority culture because of the negotiation that must occur as a minority person growing up in America with two distinct cultures. As one might imagine, identity development is very important for understanding Asian Americans, which significantly contributes to the diversity among Asian Americans. Two important factors to consider are assimilation and ethnic identity. *Assimilation* refers to the process of becoming Americanized or adopting the American culture. *Ethnic identity* refers to the retention of customs, attitudes, and beliefs of the culture of origin. These two factors represent issues that each Asian American must resolve over his or her life. For example, Bob's actions and beliefs about life may be very similar to the mainstream White European American culture, and he may dismiss the importance of his Asian heritage and have almost nothing to do with it. On the other hand, Bob's sister, Soo, may embrace both American culture and her Asian heritage and feel comfortable with members of each cultural group (bicultural).

Samuel Y. Song and Wakako Sogo

See also Acculturation; Cultural Deficit Model; Cultural Diversity; Culture; Ethnicity and Race; Multicultural Classrooms; Multicultural Education

Further Readings

Fadiman, A. (1998). *The spirit catches you and you fall down: A Hmong child, her American doctors, and the collision of two cultures.* New York: Noonday Press.

Frisby, C. L., & Reynolds, C. R. (2006). *Comprehensive handbook of multicultural school psychology.* New York: Wiley.

Kim, E. H., & Yu, E.-Y. (1997). *East to America: Korean American life stories.* New York: New Press.

Lee, C.-R. (1998). *Native speaker.* New York: Granta Books.

Lee, E. (1988). Cultural factors in working with Southeast Asian refugee adolescents. *Journal of Adolescents, 11,* 167–179.

Leonard, K. I. (1997). *The South Asian Americans.* Westport, CT: Greenwood Press.

Sue, D. W., & Sue, D. (2002). *Counseling the culturally diverse* (4th ed.). New York: Wiley.

Wu, F. H. (2002). *Yellow: Race in America beyond Black and White.* New York: Basic Books.

ASSESSMENT

Terms such as *measurement, evaluation*, and *test* are sometimes used as synonyms of the term *assessment.* Although these terms are related to assessment, their meanings and purpose are distinctly different. *Measurement* may be thought of as quantifying a performance or characteristic such as score on a test, height, or weight; *evaluation* denotes judgments about the quality or worth of something; and a *test* usually means a specific instrument (paper-and-pencil test) or set of procedures to measure knowledge, abilities, or other characteristics. *Assessment* is a general term that is used to encompass everything a teacher does to ascertain the level at which students have mastered the subject matter, can perform certain tasks, or exhibit certain behaviors.

Assessment includes the collection, analysis, and interpretation of various kinds of information useful for educational decisions. Leonard Carmichael and Bette Caldwell suggested that assessment can

produce direct benefits to students, as teachers use both formal and informal assessments to diagnose students' strengths and weaknesses. Assessment can provide information that helps teachers to identify students who need additional instruction, special services, or more advanced work. Assessment also can serve as the basis for teacher reflections on their instructional effectiveness. Based on data collected though various kinds of assessments, teachers can make instructional decisions about re-teaching a lesson or unit or moving ahead with more challenging lessons. Results of assessments can provide feedback to students to indicate areas in which performance needs improvement and areas in which performance is satisfactory.

Early roots of educational assessment can be traced to the one-room schoolhouse where students moved forward to the next level when the teacher determined they had mastered the necessary knowledge and skills. Tests were administered to ascertain progress, with no letter grades (A, B, C, etc.) reported. Teachers reported student progress orally or through their written notes. The growth of schools necessitated that students be divided into different levels, and percentages were used to report student progress and to identify those who were prepared for college. Comparisons were made among students, and those earning the highest and lowest scores were assigned grades of A and F, respectively. Most of the student scores fell in the middle range and thus were assigned a grade of C for average. Student scores close to the top, but not the highest, were assigned a grade of B, and those falling below the average level were assigned a grade of D. However, in 1912, research by D. Starch and E. C. Elliott raised questions about the use of percentages as a basis for assigning letter grades and screening for college admission. Their research showed great discrepancies in the grades assigned by teachers. This apparent lack of consistency among teachers in grading spurred teachers to use grading scales such as Excellent, Average, or Poor.

In the early 1900s, standardized tests began to be used in the schools as an outgrowth of their use to screen men for the military. Standardized tests were based on national norms. Federal legislation—most notably Public Law (P.L.) 94–142, which mandates a free and appropriate public education for individuals between the ages of 3 and 21; the Individuals with Disabilities Education Act (IDEA), which amended P.L. 94–142; and the No Child Left Behind (NCLB) Act—has brought about increased interest in assessment at the local, state, and national levels.

Behavioral Objectives

Benjamin Bloom, Professor of Education at Chicago University, working with his mentor, Ralph Tyler, believed that the importance in education was whether or not students achieved what they were studying, not how they were compared with others. Bloom believed that the environment is influential to a child's success in school. He believed that it was the educator's responsibility to create an environment to serve students at various ability levels. Bloom suggested that human learning could be classified into three major domains: cognitive, affective, and psychomotor. Bloom and a group of his colleagues began development of a system to describe behaviors in the cognitive domain in 1948 and development of a classification system for the affective domain in 1956.

Each domain comprises levels that range from the lowest level of learning to the highest. This classification of learning is commonly known as "Bloom's Taxonomy of Learning" or simply, "Bloom's Taxonomy." Classification of learning domains serves several purposes. They help to assure that teachers know the objectives of the lesson and stay focused on the intended outcomes, to direct students to the objectives of the lesson, and to serve as a basis for identifying the kinds of knowledge, skills, and attitudes that will be assessed. The knowledge, skills, and attitudes are usually stated in the form of instructional (behavioral or performance) objectives to communicate clearly the student's required performance, the conditions or circumstances under which the performance will take place, and the standard or criterion that will be used to assess whether or not the student has achieved the objective satisfactorily.

Types of Assessment

Student achievement may be assessed with standardized tests or teacher-made tests. A standardized test covers a broad knowledge base rather than knowledge associated with selected lessons or a specific unit of instruction. Uniform administration procedures are required in which the directions are standardized, and all students get the same or an equivalent test under the same conditions for the same amount of time. Interpretation of the results for a specific student is

based on that student's performance compared with the performance of other students at the same grade level. This is known as *norm-referenced testing.* Normative tables showing the relative performance of a student compared with that of other students in the norm group are usually provided by the test developers for standardized tests. Because one student is compared with other students in the norm group over a broad knowledge base, it is not sound educational practice for teachers to use standardized test scores as a basis for assigning class grades. Standardized tests are used best to show parents or school personnel how one student compares with other, similar students. For accurate and meaningful comparisons to be made, it is important that the norm group be representative of the student with whom comparisons are made.

Teacher-made tests are used to assess student achievement over a specific body of knowledge taught in the classroom. Teacher-made tests are known as *criterion-referenced tests.* The performance of one student is not compared with that of the group; instead, each student's response to a specific question is measured against a predetermined criterion. Clearly defined instructional objectives provide the foundation for criterion-referenced classroom assessment. In turn, criterion-referenced classroom assessments are used for the purpose of identifying the extent to which students have achieved the objectives.

Student achievement is assessed with a variety of tests and test item formats. Objective test items are commonly used to assess performance because they can be written to test higher as well as lower levels of learning and to test a broad range of subject matter. Objective tests are easily administered and scored. Scoring is more objective than in most other kinds of test items; consequently, fewer scoring errors are made. Objective test items require that students write or select a response. Written responses may be short answer or completion. Commonly used selected response items include multiple-choice, true/false or alternative response, and matching. Although objective tests remain a favorite among teachers, alternative methods of assessment are gaining in popularity. Performance ratings, essays and products, portfolios, and service learning are popular forms of alternative assessments.

Some instructional objectives require that students engage in an activity to demonstrate their level of skill. For example, athletic or musical skills cannot be assessed appropriately with a paper-and-pencil test.

The teacher will want to assess students' speed, timing, and precision as they demonstrate cognitive, psychomotor, and affective behaviors. Sequence of events is an important consideration when assessing an individual's ability to operate a motor vehicle, and quality of presentation is important when assessing the overall effectiveness of an individual's oral report delivery. Performance assessments are time consuming to administer because only one student can be assessed at a time. Also, the teacher must be keenly attentive to the performance in order not to miss important but rapidly performed steps. Rubrics such as checklists and rating scales are used to ascertain the effectiveness of the performance.

Some instructional objectives require students to construct individual and unique responses to demonstrate their competence. Responses may be in the form of a written essay or a product. Essays and products typically incorporate higher-order thinking skills in analysis, synthesis, and evaluation. A written essay, for example, may require that students critique a procedure, event, literary work, or political or philosophical view. Creating a product requires planning, organizing, and actually producing a tangible result such as a poem, play, blueprint, computer program, or birdhouse. Essays and products are time intensive for teachers to plan and score because of the individual nature of the products. Responses must be scored by an expert teacher who can ascertain whether or not responses are accurate and acceptable.

Portfolios are an excellent way to document and communicate student progress over time. Samples or exhibits of a student's initial work are usually included in the portfolio, and other exhibits are added throughout the grading period or the course. Exhibits may include accomplishments and deficiencies that are not directly related to instructional goals. For example, exhibits may include selected test scores, notes related to attitudes and behaviors that are not evident on report cards, social accomplishments, special projects, interest inventories, and interviews. Exhibits included in the portfolio are often used to supplement other kinds of assessment.

Service learning is gaining in popularity as an alternative teaching and assessment tool because it requires students to demonstrate their knowledge, skills, and attitudes in real-life situations. Students are involved in activities designed to meet instructional goals while providing a useful service to the community. The purpose of service learning is to improve

academic learning while enhancing personal skills and civic responsibility through structured projects that serve a real community need.

Item Analysis

An *item analysis* provides information about the difficulty of a test and the extent to which a specific test item discriminates among the test-takers in the same way that the overall test discriminates. Item analysis may be performed by a computer program or hand calculated by classroom teachers.

Item Difficulty

An item difficulty index is a statistical procedure that indicates the proportion of students who responded correctly to an item. The closer the index is to 1, the higher the proportion of students who responded correctly to that item and the easier the item. The term *item difficulty index* is referred to in many sources as the *item easiness index* because it refers to the percentage of students responding correctly to a specific item. Calculating this index requires that the test items for the high scoring students and low scoring students be separated from those whose scores fell in the middle range. Depending on the total number of students, this may mean using the top 25% and the bottom 25% in the analysis. If the total number of students is small, then all students may be used in the analysis by grouping students in either the high scoring or the low scoring group. The difficulty index is then calculated by adding the number in the high scoring group who got the item correct to the number in the low scoring group who got the number correct and dividing the sum by the total number of students.

Item Discrimination

Item discrimination is a statistical procedure that indicates the extent to which an item discriminates between high scoring and low scoring students on a specific test item in the same way that the total test discriminates between those who know the content and those who do not. This procedure also requires that the test items for high scoring students and low scoring students be put into separate groups. If an item discriminates adequately between the high scoring group and the low scoring group, students in the high scoring group should get a specific item correct more often than students in the low scoring group. A negative (and undesirable) discrimination would result if students in the high scoring group missed an item more often than students in the low scoring group. An item discrimination index may be calculated by subtracting the number of correct responses on one item of students in the low scoring group from the number of correct responses of students on the same item in the high scoring group and dividing this difference by the number of students in one of the groups. The resulting discrimination index will be between 0, indicating no discrimination, and 1, indicating perfect discrimination between high and low scoring students on an item. An index close to 1 indicates high discrimination between the two groups on a specific item, and thus, this is the more desirable item.

An item analysis provides insight into the way specific items function within a test. Such an analysis can help teachers to increase the reliability and validity of their tests. In addition, the analysis can provide teachers with information regarding items that are hardest for students, which are easiest, and where adjustments in content or the teaching process need to be made. In their book *Instructors and Their Jobs*, W. R. Miller and M. F. Miller illustrate statistical procedures useful for teachers conducting an item analysis.

Test Reliability

Reliability and validity are two important concepts in assessment. *Reliability* refers to the consistency or stability of a test in producing the same or similar scores over repeated administrations of the test. For example, if a student scores high or low on a test of word comprehension in the morning, then one should expect that if the same or a similar test were administered in the afternoon, the student would have the same or a similar score providing that no changes were made in the student's level of knowledge or administration of the test to bias the results. If the two administrations of the test yielded the same or similar results, then the test would be reliable; however, if there were a great discrepancy between the two administrations, then the test would be unreliable. Reliability is affected by different sources of error, which are called random errors. Variations due to individual attributes such as general knowledge, ability, skills, motivation, health, and attention are all sources of error. In addition, random errors may be due to variations in the characteristics of the test.

A test that is too short gives an advantage to students who happen to know the correct answer to a few questions, whereas students who actually have a broader knowledge base may not know the answers to the few questions asked. It is valuable for educators to know the amount of true variance (spread of a set of scores) and error variance reflected in a particular test score. In educational testing, it is assumed that every observed test score has a true score component and an error score component; the variance of the observed score is equal to the variance of the true score plus the variance of the error. However, in practice, a true test score is impossible to calculate because one would have to administer a test an infinite number of times. If this were possible, the sum of the errors and the mean of the errors would be 0. Theoretically, the true score for one individual is the average score obtained from repeated measures; it is considered to be somewhere between plus and minus a certain margin of error, called the standard error of measurement (SEM). The SEM is derived from the standard deviation of the sampling errors. This concept is illustrated by a student who scores an 80 on a mathematics achievement test and the SEM is 3.2. The student's true score is said to fall somewhere between 76.8 and 83.2. A common practice to determine stability of test scores is to compute a coefficient of correlation, known as a reliability coefficient, between two independent administrations of the same test to a large sample within a reasonable time period between administrations (test–retest reliability coefficient or coefficient of stability). Other procedures to ascertain reliability are (a) correlating the scores of alternate forms of the same test administered in succession to a large group (coefficient of equivalence) and (b) correlating the scores on equivalent forms of a test given in two independent administrations over a short period of time (coefficient of stability and equivalence). The reliability coefficient ranges from 0, indicating no reliability of the test, to 1, indicating perfect reliability.

Different reliability procedures are required to ascertain the extent to which all items in a test are measuring the same information. This is known as the internal consistency of a test. The split-half technique is a commonly used procedure to determine internal consistency. This involves administering a test, dividing the items independently into two equivalent halves, and computing the correlation coefficient between the two halves. The Spearman-Brown statistical procedure is used to calculate internal consistency using split-halves. Kuder-Richardson formulas 20 and 21 and Cronbach alpha are useful for assessing the homogeneity of items within a test. These latter procedures do not require splitting the test; rather, they are based on properties of the entire test.

Developing reliable tests is a challenge, and obtaining perfect reliability is impossible or difficult at best. Reliability is affected by the difficulty level of the test. A test that is too difficult or too easy can be reliable in that similar outcomes result. However, in the case of a test that is too difficult, the test provides chance results, and in the case of a test that is too easy, the test would not yield meaningful results. However, teachers can improve reliability by (a) creating a positive test environment, (b) allowing enough time for students to complete the test, (c) selecting or constructing tests that cover a representative sample of the content, including enough items to test the breadth of the content, and (d) including test items that discourage guessing so that students' scores remain stable from one administration of a test to another or from one form of a test to another. Reliability is concerned with the stability of test scores over time, and even though classroom teachers do not generally calculate the reliability of their tests, test developers are expected to assess and report the reliability for all standardized tests. Reliability is a necessary requirement for validity.

Test Validity

Validity is the extent to which a test measures what it was designed to measure. This means that tests are designed for specific purposes, and each test must have its own validity for the purpose for which it was designed. A valid test must also be reliable; however, the converse is not true. Reliable tests are not necessarily valid. That is, a test may consistently measure the wrong thing. Establishing test validity is thought to be a more complex process than establishing test reliability because establishing validity depends on the judgments to be made based on test results and how the results will be used. It is necessary to collect information as evidence that a test provides a true measure of such abstractions. To validate that tests provide true measures, certain information or evidence must be collected depending on the type of validity to be determined. Three common types of validity are content validity, criterion-related validity, and construct validity.

Content Validity

Assessing an individual on the complete domain of a knowledge base related to a specific subject is not usually feasible. Therefore, the typical procedure is to assess only a sample of the knowledge base. For example, teachers do not purport to ask students every possible question about a topic on a classroom test. They select a sample of test questions from which to make inferences about the students' grasp of the subject matter. Content validity is concerned with the extent to which the sample of questions is representative of the knowledge base being tested. If experts in the content area to be tested agree that the test questions are a representative sample of the knowledge base to be tested, the test is said to have face validity. Logical validity is closely related to face validity, except in establishing logical validity, all the behaviors to be measured within a specific knowledge base need to be identified and defined. Then, test questions are designed purposely to test the desired knowledge base. Content validity is established through rather subjective techniques; consequently, the potential for error is greater than in criterion-related or construct-related validity. Also, it is important that users of a test, especially achievement tests, understand the knowledge base that the test was designed to assess. Nonetheless, content validity is important to the development of all tests, and other kinds of validity are established more confidently when a test has content validity.

Criterion-Related Validity

The extent to which scores on a test correlate with an independent external variable provides evidence related to criterion-related validity. The independent external variable is the criterion that is thought to be a direct measure of the attribute of interest. For example, college admission tests are thought to be valid predictors of success in school as denoted by grade point average. The test to be validated is administered and scores are held until the criterion scores are available. Next, the scores on the college admissions test (predictor) are correlated with grade point average (criterion). The greater the correlation is between scores on the admissions test and the grade point average, the stronger the criterion-related validity is. This is referred to as predictive validity. The more accurately a test predicts is directly related to the strength of the correlation between the scores on the predictor test and the criterion.

Concurrent validity can be determined by correlating scores on a measuring instrument to be validated with a current score. A correlation coefficient is calculated as in predictive validity; however, the difference between predictive validity and concurrent validity is that in concurrent validity, the test to be validated is correlated with current scores on some measure. For example, the correlation of scores on a standardized mathematics test with students' current grades in mathematics provides concurrent validity. Concurrent validity is most useful for diagnostic purposes or assessment of current status.

Construct Validity

Construct validity refers to the extent to which a test measures a theoretical variable (construct) or trait. Many variables of interest in educational research are abstract. For example, intelligence, motivation, interests, self-concept, and self-efficacy are abstractions for which researchers must determine appropriate indicators that will give an assessment of these attributes. Establishing construct validity is an ongoing process, and several different procedures may be used for this process. Empirical data can be used to indicate the extent to which items on the test represent the construct. Positive correlations among items on the test and positive correlations between scores on the test and other valid observations indicate that the items include the elements predicted by the construct. Extremely low or negative correlations are indicators that the test items do not represent the construct or that an inappropriate construct is being measured. In addition, visual inspections of the items and questioning the intent of each item can help to ascertain the extent to which the items represent the construct.

Marie Kraska

See also Evaluation; Reliability; Testing; Validity

Further Readings

Allen, M. J., & Yen, W. M. (2002). *Introduction to measurement theory.* Prospect Heights, IL: Waveland Press.

Anderson, L. W. (2005). *Classroom assessment: Enhancing the quality of teacher decision making.* Mahwah, NJ: Lawrence Erlbaum.

Butler, S. M., & McMunn, N. D. (2006). *A teacher's guide to classroom assessment: Understanding and using*

assessment to improve student learning. San Francisco: Jossey-Bass.

Haladyna, T. M. (2002). *Essentials of standardized achievement testing: Validity and accountability.* Boston: Allyn & Bacon.

Miller, W. R., & Miller, M. F. (2002). *Instructors and their jobs* (3rd ed.). Homewood, IL: American Technical Publishers.

Salkind, N. J. (2006). *Tests & measurement for people who (think they) hate tests & measurement.* Thousand Oaks, CA: Sage.

Tanner, D. E. (2001). *Assessing academic achievement.* Needham Heights, MA: Allyn & Bacon.

Thorndike, R. M. (2005). *Measurement and evaluation in psychology and education* (7th ed.). Upper Saddle River, NJ: Prentice Hall.

ASSISTIVE TECHNOLOGY

As technology is reinvented to make everyday functions easy, its marriage with assistive devices significantly affects the life of individuals with disabilities. This technology is referred to as *assistive technology (AT).* Assistive technology, including assistive, adaptive, and rehabilitative devices, aids individuals with disabilities in achieving greater independence and self-confidence in their daily lives; specifically, AT enables individuals with a range of cognitive, physical, or sensory impairments to have alternative ways of performing and participating in society. AT also assists in communication, information processing, education, work, and recreation activities.

The importance of AT in the lives of people with disabilities was pointed out in the Individuals with Disabilities Education Act (IDEA; also called the Tech Act) as "any item, piece of equipment, or product system, whether acquired commercially off the shelf, modified, or customized, that is used to increase, maintain, or improve the functional capabilities of a child with a disability" [P.L. 108–446, Title 1, Part A, § 602(1)]. According to this definition, AT encompasses a wide range of devices from low-tech supportive tools such as battery interrupters to high-tech advanced and customized electronic equipment such as computers. AT can include mental aids, communication aids, alternative computer access, visual aids, or aids to augment hearing to increase the involvement of persons with disabilities in learning activities.

Through IDEA, AT can be utilized in educational settings with a new assigned term, *special education technology,* to provide students who require a differentiated instructional treatment with increased and independent learning opportunities. Appropriate AT services can also be offered in free public education.

IDEA highlights AT services, aiming to enable students to have full inclusion and least restrictive environment in schools. This law defines AT service as "any service that directly assists an individual with a disability in the selection, acquisition, or use of an assistive technology device" [P.L. 108-446, Title 1, Part A, §602(2)]. This service includes the following:

- The evaluation of the technology needs of the individual with disability, including a functional evaluation of the individual in the individual's customary environment
- Purchasing, leasing, or otherwise providing for the acquisition of AT devices by such individual
- Selecting, designing, fitting, customizing, adapting, applying, maintaining, repairing, or replacing of AT devices
- Coordinating and using other therapies, interventions, or services with AT devices, such as those associated with existing education and rehabilitation plans and programs
- Training or technical assistance with AT for such individual, or, when appropriate, the family of such individual
- Training or technical assistance for professionals including those individuals providing education and rehabilitation services

The intention of IDEA is that AT devices and services must be provided at no cost to the student, student's parents, and the school staff as part of the student's special education, related services, or supplementary aids and services to facilitate the student's education. However, the school may use whatever state, local, federal, and private resources are available.

Advocates of AT list the individuals who might benefit from special education technology:

- Individuals with mental or physical impairments that interfere with learning or other life functions
- Individuals with mild learning or cognitive problems like learning disabilities or cognitive disabilities ranging from mild to severe that interfere with learning or other life functions

Because AT has demonstrated an ability to provide individuals with disabilities opportunities that fit into their individual needs and abilities, the appropriate

use of AT may be delineated when it enables the individual to do the following activities:

- Perform functions that cannot be achieved by other methods
- Approximate normal fluency, rate, or standards—a level of accomplishment which could not be achieved by any other means
- Participate in programs or activities which otherwise would be closed to the individual
- Complete tasks that otherwise are too laborious to be attempted on a routine basis
- Concentrate on learning or employment tasks rather than mechanical tasks
- Have greater access to information
- Have normal social interactions with peers and adults

There are approximately 20,000 AT devices currently available. These can be located through online resources such as Abledata, which is supported by the U.S. Department of Education. Because of the large number of available AT devices, specialists and professionals must give thoughtful consideration to the selection and procurement of appropriate devices for impaired students. AT devices and applications include various types of accommodations and adaptations that enable individuals with disabilities to function more independently and confidently. The following section lists and categorizes common AT devices and applications based on individuals' impairments.

Assistive Technology Devices

Communication

In this category are devices or equipment designed to help individuals with speech disabilities or writing difficulties (e.g., no or very little verbal skills, limited language proficiency, etc.) to communicate effectively and to have alternate methods of communicating needs, feelings, ideas, and perceptions. Speech and augmentative communication devices include on-screen communication boards, auditory or visual scanning, speech synthesizers, text-to-speech software and hardware, head wands, light pointers, mouth sticks, signal systems, telephony equipment, and talking word processing with writing support. Writing and typing devices include tactile devices, Braille devices, note-taking devices, spelling devices, word prediction/completion software, modified typewriters, portable typewriters, and electronic and software dictionaries.

Access

These hardware and software products enable individuals with barriers to access to interact with learning tools and engage in classroom or home activities. *Alternative input devices* include alternative and adaptive keyboards, expanded keyboards, keyguards, alternative and ergonomic mouse/pointing systems, head-operated pointing devices, eye-gaze pointing devices, mouth/tongue pointing devices, Morse code input devices, brain-actuated pointing devices, switches, touch screens, voice input systems, speech-to-text software, voice recognition/voice command software, dictation software, on-screen keyboards, cursor enlargement software, ergonomic computer-based equipment, scanners and optical character recognition, joysticks, and trackballs. *Processing devices* include abbreviation/expansion and macro programs, access utilities, menu management programs, reading comprehension programs, writing composition programs, and writing enhancement tools. *Alternative output devices* include Braille display/output devices, Braille embosser/printers, screen-reading software, screen magnification/enlargement software, large print monitor, and speech synthesizers. *Accessible software* includes software applications adapted for individuals with disabilities, operating system accessibility options, and accessible web browsers. *Universal design* (also called *design-for-all* or *accessible design*) includes design methods, techniques, and guidelines for making computers and their applications fully accessible to individuals with disabilities.

Learning and Studying

These devices help individuals with high-incidence disabilities (learning, behavior, or cognitive disabilities) to increase, maintain, or improve their functional capabilities. Such devices include Post-It™ notes, picture or written schedules, social stories, written or picture-supported directions, and editing devices (e.g., correction fluid, correction tape, correction pen, highlight tape, etc.), sentence windows, graphic organizers, single-word scanners (reading pens) or handheld scanners, portable or talking word processors, handheld computers, voice-recognition products, software for organizing ideas and studying, electronic organizers or reminders, word-prediction software, talking electronic device or software to pronounce challenging words, graphic organizer software, software for

concept development, manipulation of objects, math computations, portable word processor to keyboard instead of write, closed-captioning television, text-reading software, and tactile or voice-output measuring devices.

Vision

Products designed to assist the blind and visually impaired individuals include auditory and speech output devices, reading machines, scanning/document reading systems, optical character recognition (OCR) systems, electronic book readers, talking equipment (clocks/watches, calculators, etc.), Braille devices, Braille transcription and translation devices, screen magnifier/enlarger, closed circuit television for magnifying documents, book holders, manual and electric page turners, large button phones, speaker phones, large-print books, taped/audio books, light boxes, high contrast materials, thermoform graphics, synthesizers, scanners, and descriptive video services.

Hearing

Products designed to assist individuals who have deaf and hearing impairments and/or auditory processing problems include assistive listening devices, hearing aids, infrared/personal amplification systems, audio/FM loop systems, FM amplification systems, TV amplifiers, TV decoders, visual signaling and alerting systems, tactile alerting systems, telephony and accessories, text telephones, telecommunication devices for the deaf (TDD), teletypes (TTY), adapted phones, real-time captioning, telecaption decoders, and cochlear implants.

Daily Living

Self-help devices that assist persons with disabilities in daily living activities, such as dressing, personal hygiene, bathing, home maintenance, cooking, and eating, include reaching devices, adaptive clothing, modified eating utensils (e.g., specialized spoons for self-feeding), feeding accessories and devices, adapted books, pencil holders, time management aids, dressing aids, adapted personal hygiene aids, bathing accessories, grab bars/grips/handles, incontinent supplies, mechanical transfer lift, shower/bath chair, toileting accessories, transfer board, wheeled bath chair/commode, and bathtub seats.

Environment

The use of these aids and equipment help to remove or reduce physical barriers for individuals with disabilities in educational and workplace settings. *Environmental controls and switches* are primarily electronic systems that enable someone with limited mobility to control various appliances, lights, telephone and security systems in their room, home, or other surroundings. These systems include environmental control units, electronic appliance switches, switch mounting systems, home automation systems, signaling and alerting devices, home alarms, television adaptations, smoke alarms, and telephone ringers. *Home-workplace adaptations* include worksite/school/home design or modification for accessibility, architectural accommodations, structural adaptations, building/home ramps, home elevators, wheelchair lifts, pool lifts, bathroom changes, automatic door openers, expanded doorways, adapted furniture, adapted doorknobs, alternative doorbells, lowered counters, and specially designed bath areas.

Ergonomic Equipment

Ergonomic equipment is low-tech assistive equipment or devices designed to reduce the likelihood of repetitive stress injuries often associated with work-related situations. These devices include adjustable workstations, industrial workstations, office workstations, adapted furniture, writing aids, modified seating and lighting, arm/wrist supports, and back supports.

Mobility and Transportation

These products help mobility-impaired persons move from one place to another and give them independence in personal transportation. Such products include standing/walking aids, transfer aids, stair lifts, walkers, scooters, wheelchairs and three-wheeled chairs, adapted bikes and tricycles, car seats, beds, stretchers, patient chairs, ramps, recliners, strollers, travel chairs, wheelchair trays, driving controls, seat belts, vehicle conversions, patient and wheelchair lifts, wheelchair loaders/carriers, and wheelchair restraint systems. Ambulatory aids include canes, cane accessories, crutches, walkers, and walker accessories. Vehicle conversions include car-top carriers, custom cars and vans, adaptive driving controls, hand-controls, child restraint systems, ramps, and lifts.

Prosthetics and Orthotics

Prosthetics and orthotics serve to replace, substitute, or augment missing or dysfunctioning body parts with artificial limbs or other aids. These devices include splints, braces, foot orthoses, helmets, restraints, and supports.

Recreation and Leisure

These products help persons with disabilities to participate in sports, social, and cultural events and interact with their peers without disabilities, which allow disabled people to focus less on individual differences. Such products include adaptive sports equipment for skiing, biking, swimming, running, boating, adaptive music with symbols, adaptive controls for video and table games, adaptive fishing rods, cuffs for grasping paddles or racquets, and seating systems for boats.

Seating and Positioning

These products assist impaired individuals in maintaining body alignment, upright posture and trunk/head support, and reducing pressure to the skin. Such products include adapted and modular seating, cushions and wedges, contour seats, lumbar support seats, standing tables, positioning belts, braces, wheelchair modifications and cushions, seat lifts, bolster chairs, corner chairs, therapeutic seats, postural support hardware, postural support systems, and pressure monitors.

Tufan Adiguzel and Kimberly J. Vannest

See also Disabilities; Educational Technology; Individuals with Disabilities Education Act; Least Restrictive Placement; Special Education

Further Readings

Assistive Technology Act of 2004, P.L. 108–364 (2005). Washington, DC: U.S. Government Printing Office.

Edyburn, D. L. (2004). Rethinking assistive technology. *Special Education Technology Practice*, 5(4), 16–23.

Hasselbring, T. S., & Glaser, C. H. W. (2000). The future of children. *Children and Computer Technology*, 10(2), 102–122.

Individuals with Disabilities Education Improvement Act of 2004, P.L. No. 108–446 (2005). Washington, DC: U.S. Government Printing Office.

RehabTool. (2007). *RehabTool: Deliver assistive technology products and services.* Retrieved January 10, 2007, from http://www.rehabtool.com

ATHLETICS

College athletics in the United States has prospered significantly since its inception. It has also been faced with controversy ranging from threatening the academic integrity of higher education to extreme levels of commercialization. Regardless, college athletics has provided millions of students with opportunities to not only compete in athletics but also achieve an education and ultimately a college degree. This entry examines college athletics in higher education. It uses both historical and sociological sensitivities to provide a conceptual framework for this analysis. These sensitivities provide a historical context to understand the origin and development of college athletics and a sociological context to understand how college athletics has evolved into a social spectacle that provides millions with a source of entertainment and a cultural event that inform their daily life.

Origin and Governance of Men's and Women's College Athletics

College athletics has a very modest beginning. The origin of college athletics dates back to the Gilded Age during 1865–1900. It began as student-controlled sporting opportunities for members of the university student body, and because the sports were not sanctioned by the institution, they did receive financial support. As these activities became popular and profitable enterprises, alumni, faculty, and administrators took control, which created a shift from college athletics being informal, student-controlled activities to formal, university-sanctioned events; this move from informal to formal required additional administrative and governance structure to oversee these events.

A major part of college athletics history is the development of its organizational and governance structures. Therefore, a major transition to denote a shift in college athletics took place in 1905 when President Theodore Roosevelt summoned 13 institutions to the White House to address the need for athletic reform, more specifically, the need for rule changes in the game of football. Thus, the Intercollegiate Athletic Association of

the United States (IAAUS) was formed by 62 institutions to address additional concerns that emerged since the inception of college athletics. In 1910, the IAAUS changed its name to the National Collegiate Athletic Association (NCAA) and continued to grow beyond its initial origins of being an advising and rule-making association and sought to address several issues that were challenging the very nature of college athletics.

For example, due to the increase in popularity and profit of college athletics, issues regarding academic integrity, professionalization, institutional control, and the safety of players were major concerns that created a need for athletic governance. College athletes at several major schools, at this time, were basically recruited to play sports; thus, the academic integrity and mission of the institution was being challenged. The increased popularity also gave athletic departments unlimited freedom in operating beyond the educational philosophy of the university, which required a need for greater institutional control where all departments associated with the university would ultimately report to the university's president and comply with the educational mission of the institution. Finally, several deaths occurred during football competition due to game rules and equipment.

College athletics faced growing pains due to expansion in membership and an increase in postseason championship games. It was not until after World War II that issues concerning recruiting and financial aid were addressed whereby standards were being established to minimize athletic deviance. However, the NCAA continued to face challenges that required professional governance and a shift from part-time leadership to full-time leadership. In 1951, Walter Byers became executive director and began to chart a new course for the member institutions, in which the NCAA, on behalf of its member institutions, would gain control over live television coverage of football games and postseason bowl games.

Governance for women's athletics had a different philosophy than that of men's. Although women's athletics has been around as long as men's athletics, it did not receive equal support, and women did not have a wide variety of sports to compete in at the collegiate level. Both men's and women's athletics had strong connections with university physical education departments; however, women's sports governance adhered to more a process orientation rather than outcome orientation. In other words, competing in sports was about the process of learning desirable skills and abilities and not solely about winning. Women's athletics was less commercial and did not initially offer scholarships or recruit outside of the respective institutions' student bodies.

The organization of women's athletics on the national level did not take place until 1941. Its first national championship game was in the sport of golf. As women's athletics began to increase and colleges across the nation began organizing women's sports teams, the need for additional governance increased. It was during the late 1950s and early 1960s that a committee with representatives from three organizations (the National Association for Physical Education for College Women, the National Association for Girls' and Women's Sports, and the American Federation of College Women) consolidated its forces under one organization called the Division for Girls' and Women's Sports, and from this consolidation the National Joint Committee on Extramural Sports for College Women was formed, with the primary responsibility governing women's intercollegiate athletic programs. From this committee, the Commission on Intercollegiate Athletics for Women (CIAW) was formed in 1967 with two major purposes: (1) to provide a structure and governance of women's collegiate athletics, and (2) to oversee and sponsor national championships for women collegiate sports. However, in 1971, the Association for Intercollegiate Athletics for Women (AIAW) evolved from the CIAW and became the governing body for women's athletics.

Throughout the various levels of consolidating and organizing, women's sports continued to grow in membership, to around 280 schools, and in championship games played at the national level. What assisted in the growth and support of sporting opportunities for women and girls, especially women at the collegiate level, was the passing of Title IX of the Education Amendments in 1972 by the Congress of the United States. Title IX was enacted to prevent federally funded institutions from denying support and benefits and/or discriminating on the basis of sex. Although Title IX applied to all educational activities, it had significant impact on the increase in girls' high school and women's collegiate sports. To ensure its proper implementation in the area of sport, a three-prong test was developed to inform institutions of their compliance. The following questions were posed: (1) Are athletic opportunities substantially proportionate to the student enrollment? (2) Is there evidence of consistent growth of athletic opportunities for the previously marginalized gender? (3) Is there

complete and effective accommodation of the interest and ability of the previously marginalized gender?

Ultimately, despite the fact that the AIAW had around 1,000 members during its governance of women collegiate athletics, when the NCAA decided to offer women's championships, it threatened the AIAW's existence. When the NCAA began to offer incentives packages to women's teams participating in championship games, the AIAW began to lose its status and popularity. One major misfortune to the AIAW was when the National Broadcasting Corporation canceled its contract with the AIAW, causing many schools to participate in the NCAA's national tournament instead. After failed attempts to sue the NCAA for its attempts to monopolize collegiate sports, the AIAW ceased to operate in 1982.

With the merger of women's and men's collegiate sports under the governance of the NCAA, athletic departments witnessed an increase in scholarships and larger recruiting and operating budgets. Furthermore, despite the growth of both women's and men's collegiate athletics, several problems emerged from this merger regarding equitable treatment and support for women's athletics. Thus, there has been a continual need for women's sports to seek to enforce the guidelines of Title IX, which is also referred to as gender equity, to ensure equitable treatment.

The origin and governance of collegiate athletics has been a long and winding road to its current destination at this point in history. From its humble beginnings, collegiate athletics has grown into a multibillion-dollar industry with multiple streams of income deriving from the sales of broadcasting rights; sponsorship; collegiate licensing merchandise; naming rights of stadiums, arenas, and championship events (including bowl games); and more.

The Role of the Athlete in Higher Education

A key area to address in the growth of college athletics is the role of the athlete at the collegiate level, which has changed considerably since the first intercollegiate athletic competition. Initiated originally as extracurricular or intramural activities, some of these competitions have evolved into multimillion-dollar spectacles. Once events intended to break the monotony of the rigorous academic pursuit, these competitions, especially in basketball and football, have become main staples and major attractions at many collegiate institutions in the United States. Thus, the athlete has evolved from simply a leisurely participant in these competitive engagements to an athletic celebrity on campuses throughout the United States.

To further distinguish the role of the athlete in college athletics, it is important to examine the major divisions of competition within the NCAA, as it comprises 1,025 members at different divisions and subdivisions. Another, smaller collegiate athletic governing body is the National Association of Intercollegiate Athletics, which has about 282 members and is considered more amateur than the NCAA in its role and philosophy. The major level of athletic competition is Division I-A, which comprises 119 members. Some of the basic requirements for institutions to hold membership at this level are that they must fund at least seven men's and women's sports or six men's and eight women's sports; they must also have at least two team sports for women, and each playing season must have a fair gender representation. Another basic requirement is related to the scheduling of events and the participant minimum for each sport.

These NCAA Division I institutions are the top-tier athletic programs in the country, and the athletic operating budgets of the top 20 of these institutions range from $32 to $75 million. They sponsor football programs with stadiums that seat 60,000 to 100,000 people and basketball arenas that range in seating from 10,000 to over 25,000 people. These programs operate like corporations, where athletic talent is the major commodity. These programs require a premium on athletic talent; thus, they spend a significant amount of time and money recruiting the best athletes. This form of corporate athleticism or athletic capitalism places the athlete within controversial and sometimes diametrically opposing roles, especially in revenue-generating sports (i.e., football and basketball). High-level athletic demands often place academic demands as a lower priority. Being both a student and an athlete creates challenging roles that require a healthy balance that many young men and women competing in intercollegiate athletics learn to master effectively.

Graduation rates are often the determining factor to whether athletes are successfully balancing the role of student and athlete. At this level, the NCAA reported in 2006 that athletes graduated at a rate of 77%, whereas the regular student body graduated at a rate of 61%. The revenue-generating sports of football and men's basketball have had the lowest graduating rates, at 55% and 45%, respectively. Black male athletes

were reported as graduating at 49% in football and 38% in basketball. Women, especially White women athletes, and White male athletes fare better than their Black male counterparts at balancing the role of student and athlete and continuing on to graduation.

One of the factors affecting graduation rates for athletes in the sports of football and basketball is early departure into professional sports. These early departures count against the institution, thus lowering its graduation percentage rate. Despite the age limit restriction instituted by the National Basketball Association and the National Football League's restriction on not allowing underclassmen the opportunity to compete until a minimum of 3 years after their high school graduation, many athletes have turned professional without the reward of graduating with a degree. Intercollegiate basketball has been a victim of athletes spending only one or two semesters on campus before entering the draft. This gives them enough time to meet the minimum age requirement but unfortunately alters graduation rate calculations.

One of the remedies for the low graduation rates among athletes in sports that generate millions of dollars a year in revenue is the Academic Progress Report (APR). The APR was designed to assist in reforming college athletics by improving academic success and graduation rates among athletes. It is calculated whereby one point is given to a team each term a scholarship athlete meets the required academic eligibility standards; an additional point is given if the athlete remains with the institution. A cutoff score of 925 corresponds to graduating 59% of scholarship athletes. If a team does not meet the cutoff on what the NCAA refers to as a basis that is statistically significant, the university can be penalized by losing a scholarship in that sport.

The need for athletic reform speaks to the demand athletes have had in balancing the roles of student and athlete. It also speaks to the incongruence many critics see between the earning of millions of dollars off the blood, sweat, and tears of young vulnerable men and women and the lack of academic integrity demonstrated each year when institutions are exposed for academic fraud. Athletics at other divisions have not suffered the same criticism because of fewer demands to generate revenue.

The next collegiate levels of athletic competition are the Division I-AA and I-AAA levels. At the Division I-AA level, which consists of 116 members, the athletic programs are considerably smaller. There are

91 members at the Division I-AAA, and these institutions do not have football programs. However, they compete in a variety of other men's and women's sports.

The next level within the governance of the NCAA is Division II, which comprises 281 members. These members must fund a minimum of five sports for men and five for women or four for men and six for women. Similar to Division I members, Division II members must also have a minimum of two team sports for each gender and have fair gender representation for each playing season. One of the major differences between Division I and II is that financial support for students at Division II schools is generally a combination of money from scholarships, grants, and student loans, whereas Division I schools are able to provide full athletic scholarships to more athletes. Therefore, the range of students being recruited is limited to state and local athletes for most Division II schools, whereas the range for recruitment at Division I schools is greater— including national and international recruitment of athletes. The role of the Division II athlete is less labor intensive than the role of the Division I athlete.

The final level to discuss is Division III member institutions. There are currently 420 institutions that are members at this level, which makes it the largest division governed by the NCAA. These institutions are relatively smaller in student body population than the other two divisions: 500 to over 10,000. One of the differences between Division III and Divisions I and II is that Division III member institutions do not offer any athletic-related financial aid. At this level, the sports are considered extracurricular activities for the student body population; thus, they do not generate revenue. Athletic participation at this level is considered purely amateurism by many critics of college athletics. Because athletes do not receive any financial aid or endowments, they are basically pursuing athletics for no financial benefit. Thus, academics are generally the primary focus of the student, and athletics is secondary. The major reason is that very few athletes become professional athletes from this division. Therefore, the philosophy of these programs is more process oriented instead of outcome or profit driven; it is for the love of the game.

Future Directions

College athletics has evolved into a profitable industry that has combined the pursuit of athletic excellence with the academic mission of institutions of higher learning.

At the NCAA Division I level, the challenge has been in operating revenue-generating activities within a non-profit entity. This has drawn considerable criticism from scholars and other critics proclaiming that the academic integrity of the institution is threatened, especially when scandals of academic deviance and abuse incurred by athletes and athletic administrators tarnish the images of institutions of higher education.

Regardless of these challenges, the momentum for college athletics continues to (a) supply the student body with an avenue to unite and share in a common tradition and sociocultural practice; (b) provide athletes with the opportunity to use their athletic talents in exchange for an athletic scholarship (especially at the NCAA Division I and II levels) and, for many, a college education; (c) provide athletes with the opportunity to develop transferable and marketable skills (e.g., character development, hard-work ethic, ability to work with members of a team to achieve a common goal) within the highly competitive environment of college athletics; and (d) give national and international exposure to institutions that previously have not enjoyed this privilege. Thus, the future of college athletics presents an ambiguous forecast: continual controversy and criticism at the Division I level, if athletic reform is not successful in managing the discrepancies between athletic capitalism and the educational, research, and service mission of higher education; and for the smaller divisions, college athletics will continue to provide the aforementioned services and blend the pursuit of athletic excellence with the context of higher education if these smaller institutions adhere to the model of amateurism.

Billy Hawkins

See also Curriculum Development; Family Influences; Physical Development

Further Readings

2006 NCAA Division I federal graduation rate data. Retrieved on January 1, 2007, from http://web1.ncaa.org/app_data/instAggr2006/1_0.pdf

Adler, P., & Adler, P. (1991). *Backboards and blackboards: College athletes and role engulfment.* New York: Columbia University Press.

Bowen, W., Levin, S., Shulman, J., & Campbell, C. (2003). *Reclaiming the game: College sports and educational values.* Princeton, NJ: Princeton University Press.

Duderstadt, J. (2000). *Intercollegiate athletics and the American university: A university president's perspective.* Ann Arbor: University of Michigan Press.

Hawkins, B. J. (2001). *The new plantation: The internal colonization of Black student athletes.* Athens, GA: Sadiki Press.

King, C. R., & Springwood, C. (2001). *Beyond the cheers: Race as spectacle in college sport.* Albany: State University of New York Press.

Savage, H., Bentley, H., McGovern, J., & Smiley, D. (1929). *American college athletics.* New York: Carnegie Foundation for the Advancement of Teaching.

Shulman, J., & Bowen, W. (2001). *The game of life: College sports and educational values.* Princeton, NJ: Princeton University Press.

Smith, R. A. (1988). *Sports and freedom: The rise of big-time college athletics.* New York: Oxford University Press.

Sperber, M. (2000). *Beer and circus: How big-time college sports is crippling undergraduate education.* New York: Henry Holt.

Zimbalist, A. (2001). *Unpaid professionals: Commercialism and conflict in big-time college sports.* Princeton, NJ: Princeton University Press.

Web Sites

National Collegiate Athletic Association: http://www.ncaa.org

ATTACHMENT

Attachment, an important component of affective development, is the affectional bond that a child has with significant others. This entry addresses what is known about the link between attachment and education and covers three main areas. First, this entry discusses children's interest in and ability to participate in relationships that promote learning, including the relationship with the primary caregiver as well as relationships with early care providers and teachers. Next, this entry discusses specific language, cognitive, and social/emotional outcomes associated with attachment. Finally, this entry addresses biological underpinnings that begin to explain why attachment may be linked with educationally relevant outcomes in young children.

Participation in Relationships That Promote Learning

In discussing early relationships, it is important to begin with an overview of key concepts of attachment theory, as applicable to educational contexts. First, for

the purpose of research, attachment relationships are classified into distinct categories based either on a laboratory separation-reunion paradigm (Strange Situation Procedure) or observation of the child and a caregiver in naturalistic context (the Attachment Q-Set), both with strictly defined scoring systems. Broadly, attachment is viewed as secure when the child responds with warmth and trust to a particular caregiver. Attachment is viewed as insecure when the child responds to a caregiver in either cool/avoidant or overly dependent/clingy ways. Because attachment relationships are considered to be unique to each child–caregiver dyad, children are thought to have the possibility for secure relationships with the different adults in their lives, particularly during infancy but also into the school years as children meet new teachers and have the possibility to enjoy trusting relationships with them, even if they did not arrive at early care or education settings with that trust in a prior caregiver. Thus, while parent–child relationships are considered to be most salient for a child's development, a child can have unique attachment relationships with different adults and, thus, unique relationships with each teacher. The security of these relationships may be different from the child's relationship with his or her parents.

For the most part, however, children have relationships with others that are similar to the relationships with their principal attachment figures at home, mainly because children form "internal working models" about relationships based on these early experiences. According to this important attachment concept, individuals use a frame of reference from which to view new experiences. Related to learning and, in particular, the academic environment, an individual's internal working models will affect the way new information and experiences are filtered, the responses an individual evokes from others, and the niches they develop for future experiences. Thus, a child enjoying a secure relationship with the mother is seen as more likely to establish secure and nonconflictual relationships with teachers. Similarly, a child enjoying a secure relationship with the mother is seen as more likely to establish secure and nondependent relationships with teachers. Thus, internal working models are thought to create templates for the quality of future relationships with other attachment figures. Although infants, school-age children, and adolescents may evoke relationships that are similar to what they have experienced, early care professionals and educators can respond to insecure children in ways that they are not accustomed to, thereby providing the opportunity for insecure children to experience secure relationships, perhaps for the first time.

Children's interest in and ability to participate in relationships is thought to be important for early learning. According to attachment theory, processes that lead children to acquire school competencies are rooted in the stability and quality of interactions in early relationships. For example, in the context of early literacy development, researchers A. G. Bus and M. H. van Ijzendoorn have found that secure attachments to primary caregivers are linked with more optimal interactions during joint book reading with that caregiver, specifically greater frequency and better quality of interaction. When securely attached dyads are reading together during the early and preschool years, they are more likely to be attentive to the reading material and to enjoy their time in this activity. The researchers concluded that securely attached children are more likely to be growing up in an environment where they become interested in literacy, because they are interested in interactions with their caregivers.

As stated earlier, an attachment relationship can form between a child and an adult other than the primary caregiver, such as an educator. Play-based child assessment of this postulate via child drawings reveals child attachment, both secure and insecure, with educators. These relationships can play a significant role in shaping both children's motivation to learn and their classroom learning experiences.

It is important to note that educators bring to their relationships with students not only their own formal training but also their personal experiences and attachment styles. Individual educator strategies for supporting learning through secure child–educator relationships include the following: understanding of socioemotional development, promotion of a supportive socioemotional climate in the classroom, attentiveness to each child, and clear and open expectations about students' academic feedback and success. These strategies have important implications for child learning outcomes, as school-age children and adolescents remember information more readily when there is a positive valence attached to the setting.

Attachment-Relevant Qualities and Child Outcomes

Research indicates that securely attached children are more likely than insecurely attached children to show

advanced levels of symbolic play. Studies have pinpointed sensitive maternal involvement (a significant predictor of attachment) as being linked with sophisticated symbolic play in toddler-age children. More specifically, both the level of play and length of time that a child engages in symbolic play are enhanced by the mother's making suggestions and communicating with the child during the child's play session.

These advantages for securely attached children in early childhood may, in part, be related to theory of mind capacity—the ability to consider another individual's perspective. Such parents often use reasoning in conversation with their children and discuss emotions. These behaviors are known to foster early social cognition in children.

Securely attached children have been found to show a greater "secure readiness to learn." Research indicates that when primary caregivers have established secure relationships with their young children, these children are more likely to be more attentive at school and to be interested in the learning environment.

Attachment also has been shown to affect expressive language outcomes, more so than cognitive outcomes. Meta-analyses have documented that securely attached children are more competent in the language domain than are their insecurely attached counterparts. Such language competencies may be linked with the relational context in which skills are developed. It is thought that language development may best be stimulated in the context of sensitive caregiver–child interactions.

Of the several different insecure attachment types, the disorganized attachment classification is noteworthy in terms of educational outcomes. Children who show this pattern are more likely to have been exposed to a difficult, unstable, and fear-evoking home environment than other insecure children, such as abuse and neglect. It is posited that the anxiety raised in task performance related to the reactions of others, for example, educators and peers, might cause dysregulated thought processes and, in turn, affect further learning. It is also suggested that difficulty with self-regulation in this group affects learning. Overall, cognitive functioning impairments in children with disorganized attachment places them at risk for academic problems.

In addition, securely attached school-age children and adolescents have better peer relationships than do insecurely attached children. Compared with other insecure or secure children, the disorganized group is likely to show more behavior problems, including aggression.

For children with learning differences, educator–child attachment is a particularly important factor in child adjustment. Secure educator–child attachment has been shown to mediate adjustment for children with learning differences. In this way, secure relationships serve as a protective factor to support learning via better child adjustment.

Some Biological Underpinnings

The brain develops in a predictable way, from the most simplistic to the most complex systems. It is generally thought that there are windows of time in which neuronal systems develop and that environmental influences during these time periods can significantly affect the developing brain by either hindering or supporting future development. The early years of a child's life are important foundational years for neurological development because environmental impacts during this period have the potential to affect later, more complex, brain development. Security of relationships has been shown to modulate arousal and attention. The majority of empirical work done to establish this relationship has been done with insecurely attached populations.

Additionally, information is stored in the brain in a use-dependent way. Thus, the more an area of the brain is activated, such as the stress response system, the more developed it becomes. Because early experiences play such a large role in shaping the stress response system, these presentations become more ingrained when, and if, children are exposed to chronic stress.

Implications

It is important for educators to see opportunities for establishing secure attachments. Rather than viewing children as having "within the child" characteristics, it is important for educators to understand that children have the capacity for different relationships, and in these different relationships, they can manifest various individual characteristics. In more mutually rewarding relationships, children have the capacity to learn more optimally, as they attach a positive valence to the school environment. Positive and secure attachments to educators and to peers at school are part and parcel of secure attachments to school. Security of

relationships not only helps children connect with school but also supports their future learning.

Shannon Altenhofen and Zeynep Biringen

See also Attachment Disorder; Cognitive Development and School Readiness

Further Readings

Bus, A. G., & van Izjendoorn, M. H. (1992). Patterns of attachment in frequently and infrequently reading dyads. *Journal of Genetic Psychology*, *153*, 395–403.

Mashburn, A. J., & Pianta, R. C. (2006). Social relationships and school readiness. *Early Education and Development*, *17*(1), 151–176.

Perry, B. (2004). *Maltreated children: Experience, brain development, and the next generation.* New York: W. W. Norton.

Pianta, R. C., Nimetz, S. L., & Bennett, E. (1997). Mother-child relationships, teacher-child relationships, and school outcomes in preschool and kindergarten. *Early Childhood Research Quarterly*, *12*, 263–280.

Tarullo, A. R., & Gunnar, M. R. (2006). Child maltreatment and the developing HPA axis. *Hormones and Behavior*, *50*, 632–639.

ATTACHMENT DISORDER

Attachment is defined as an affectional bond that ties a child to significant individuals in his or her life and endures over time. All infants raised in homes with at least one consistent parental figure grow in their emotional connection over the course of the first year of life and thus become attached. If an infant is being raised by several individuals (e.g., mother, father, grandmother, or child care professional), the infant then is in a multiple caregiving situation and will likely form an attachment with each of these important individuals, an attachment that reflects the unique quality of each of these relationships.

Although *attachment disorder* has no clear definition, it is generally defined as a failure to develop or to show a selective attachment toward at least one individual. It is thought to be caused by abuse or neglect, lack of a consistent attachment figure (as in institutional settings), or disruptions in existing relationships (as in multiple foster placements). The sensitive interactions that occur between a young child and the caregiver is thought to create a secure attachment, and

when positive interactions are lacking but there is nonetheless a consistent caregiver, an insecure attachment is thought to form. In contrast to this normative scenario, attachment disorder is the lack of the formation of a selective attachment and the impairment of the child's ability to form a focused attachment.

The *Diagnostic and Statistical Manual of Mental Disorders* (Fourth Edition, Text Revision) (*DSM–IV–TR*) divides attachment disorder into inhibited and disinhibited subtypes and requires that the symptoms be present before the age of 5 years. Attachment disorder is referred to as reactive attachment disorder. Disinhibition includes inappropriate approach, an overfriendliness toward unfamiliar adults, and inappropriate use of physical boundaries. Inhibition refers to both a child's inability to initiate or to accept comfort and a child's general fear of and withdrawal from social contacts. Although these definitions may seem clear, recent commentaries by prominent individuals in the field indicate that there is no well-accepted assessment protocol for making the diagnosis and that investigators utilize different meanings.

Attachment Theory

John Bowlby began to formulate attachment theory (integrating numerous disciplines, including, but not limited to, cognitive science, ethology, and object relations theories), and his work was the basis for a major departure from existing psychoanalytic views, which focused more on fantasy life than on actual interactions and relationships. He focused on the role of brief, as well as long-term (temporary as well as permanent), separations of infants from their caregivers and their potential impact on defensive processes of the growing infant. A focus on childhood mourning in the event of parental loss through death was also given considerable attention in terms of children's ability to attach and love. Bowlby's work was not solely on clinical populations of patients but also on normative individuals' experiences.

In 1950, Mary Ainsworth joined Bowlby's research team and was fascinated by the terrain of research ideas offered by attachment theory. It was Ainsworth and her team who developed a laboratory-based methodology for testing Bowlby's theory—the now well-known Strange Situation Procedure. In brief, an infant and caregiver (an attachment figure) are introduced into this unfamiliar context, which involves two separations from the caregiver as well as two reunions

with the caregiver. The separations are very brief (from 30 seconds to 3 minutes, depending on the reaction and stress shown by the infant). The reunion response of the infant is thought to show the infant's view of this relationship with this particular caregiver, with happy greetings (even after potential turmoil during the separations) being indicative of a secure relationship. In contrast, avoidance, clingyness, or a mixture of such behaviors and bizarre reactions are reflective of an insecure relationship with this caregiver. While there is only one form of security, as suggested above, there are several forms of insecurity, as seen in the Strange Situation, with avoidance being referred to as insecure/avoidance, clingyness being referred to as insecure/ambivalence, and a mixture of responses including bizarre behaviors such as freezing or stilling being indicative of insecure/disorganization. It should be noted that Ainsworth and her team identified the first three, whereas Mary Main and Judith Solomon identified the disorganized pattern.

Bowlby, Ainsworth, and numerous attachment researchers since then have been fascinated by the concept of maternal sensitivity and the mounting evidence on the relation between sensitivity and a secure attachment. Attachment researchers have concluded that when a parent is sensitive to an infant's cues and communications, the infant has a greater chance of forming a secure and trusting attachment connection with that person. When a parent is inconsistent or reacts in negative and/or overly emotional ways, it is thought that the infant may begin to form an anxious, resistant form of attachment to that caregiver. Such parents may even look sensitive on occasion, termed by Zeynep Biringen as *apparent sensitivity.* When a child is constantly rebuffed and rejected by a parent, an emotionally distant relationship begins to form, termed *insecure avoidant attachment.* When a child is frightened by words or deeds, is exposed to extremes in parenting dysfunction (such as abuse), or experiences significant instability in the caregiving environment, or when the parent displays difficulties in expressing normative affect, an insecure/disorganized pattern of attachment is thought to develop. These patterns of attachment are seen in normal populations and, in many ways, are viewed as normal variants, although the secure is associated with many more positive child outcomes than are the insecure forms of attachment. Bowlby asserted, and later empirical research has corroborated, under stable life circumstances, these patterns tend to persist over time.

Although attachment theory set a firm foundation for understanding social and emotional development in infants and young children in normative settings and normative families, with the range of functioning being anywhere from secure to insecure, the theory and the research have not tackled the lack of development of a focused and selected attachment (which is the definition of attachment disorder). The reason for this lack of emphasis is that most of the research on attachment has involved a specific and focused attachment figure, who brings the infant or young child to the Strange Situation Procedure. The insecure/disorganized form of attachment comes closest to an attachment disorder. However, it should be said that not all cases of disorganized attachment involve an attachment disorder, and in fact some cases of disorganized attachment may be a transient situation for some children. Thus, there is a schism between attachment theory/research based on elegant conceptualizations and/or methodologies used to understand normative populations and the current need to better conceptualize, assess, and to research the topic of the clinical phenomenon of attachment disorder.

Behaviors Seen in Other Diagnoses

Children with attachment disorder have been described as displaying aggression, lack of impulse control, resistance to authority, manipulativeness with others, lack of conscience. Such behaviors may be seen in other disorders, such as conduct disorder, or in a comorbidity between conduct disorder and attention deficit hyperactivity disorder. The current thinking is that diagnosticians should first be sure to assess for other conditions and give the label "attachment disorder" only in the rare circumstances that it is warranted.

Treatment

There are many types of treatment designed for attachment disorder, though there is no treatment that is evidence based. Approximately 70 studies have implemented attachment-based therapies (emanating from the field of attachment research), with the primary focus being the improvement of maternal sensitivity and attachment security, as conceptualized for normative populations. Interestingly, none of these studies has been implemented on groups at-risk for an attachment disorder or those who have been diagnosed.

Holding Therapy

One treatment that has created considerable controversy is holding therapy. The theory behind holding therapy is that a child's anger and rage must be released for the child to be able to function properly in society. In addition, because such a child has not experienced an attachment and is resistant to closeness, the emotional distance needs to be confronted. Physical holding, binding, forcing touch and eye contact, pinching, knuckling, licking, and punishments regarding food and water are all part of the arsenal of techniques that are used. These efforts are used to confront the rage and anger of the child and to (hypothetically) attach the child to the care provider. Opponents of this approach (which include a substantial number of attachment and trauma researchers) describe the traumatizing nature of these treatments as well as the lack of theory or research to support such extreme interactions, some of which have caused child deaths, as rebirthing is simulated in the context of restraint.

Web sites provide information on enrolling one's child in this type of therapy (often referred to as *attachment therapy*), with additional parent training (labeled *attachment parenting*) to be conducted at home. It is important that this use of attachment parenting be differentiated from a different popular use, which supports physical and social contact to create attachments (but is not considered traumatic in any sense of the word as described here). The work on holding/rage therapies are considered by the field of attachment as unsubstantiated therapies that should not be referred to as attachment therapy.

Implications

Attachment disorder is a misunderstood diagnosis, with clear linkages to maltreatment as part of the diagnosis via the *DSM–IV–TR* (therein referred to as reactive attachment disorder) but with tenuous ties to the field of attachment research. Many state that the link with maltreatment is unfortunate because the link leads to a misdiagnosis for many maltreated children who do not meet the criteria for the diagnosis or potentially any diagnosis. Similarly, those being adopted through international adoptions also are diagnosed without sufficient evidence of the disorder.

Children should not be assumed to have the disorder just because of a particular history of maltreatment or institutional rearing. The field does not have an assessment protocol, which likely will precede the empirical research in this area. In educational and other settings, it is important for children not to be labeled with an attachment disorder just because they are adopted or just because they have experienced maltreatment or foster placements. It is the inability to form selective attachments that is the hallmark of attachment disorder. If the diagnosis is accurate, however, attachment disorder is difficult to treat, and even in the most extreme of interventions available to interventionists (which is adoption into a loving home), the child may not develop the capacity for emotional connection and attachment.

Zeynep Biringen, Taylor Grant, and Deneil Hill

See also Attachment

Further Readings

Ainsworth, M. D. S., Blehar, M. C., Waters, E., & Wall, S. (1978). *Patterns of attachment.* Hillsdale, NJ: Lawrence Erlbaum.

Ainsworth, M. D. S., & Bowlby, J. (1991). An ethological approach to personality development. *American Psychologist, 46*, 333–341.

Boris, N. W. (2003). Attachment, aggression and holding: A cautionary tale. *Attachment & Human Development, 5*, 245–247.

Bowlby, J. (1969). *Attachment and loss.* New York: Basic Books.

Bretherton, I., & Waters, E. (Eds.). (1985). Growing points of attachment theory and research. *Monographs of the Society for Research in Child Development, 50*(1–2, Serial No. 209).

De Wolff, M. S., & van Ijzendoorn, M. H. (1997). Sensitivity and attachment: A meta-analysis on parental antecedents of infant attachment. *Child Development, 68*, 571–591.

Howe, D. (2003). Attachment disorders: Disinhibited attachment behaviors and secure base distortions with special reference to adopted children. *Attachment & Human Development, 5*, 265–269.

O'Connor, T. G., & Zeanah, C. (2003). Attachment disorders: Assessment strategies and treatment approaches. *Attachment & Human Development, 5*, 223–244.

Waters, E., Hamilton, C. E., & Weinfield, N. S. (2000). The stability of attachment security from infancy to adolescence and early adulthood: General introduction. *Child Development, 71*, 678–683.

Waters, E., Kondo-Ikemura, K., Posada, G., & Richters, J. E. (1991). *Learning to love: Milestones and mechanisms.* Retrieved October 11, 2006, from http://www.psychology.sunysb.edu/attachment

ATTENTION DEFICIT HYPERACTIVITY DISORDER

Attention deficit hyperactivity disorder (ADHD) is a diverse behavioral syndrome affecting 3% to 7% of children in the United States, characterized by inattention, overactivity, and impulse control problems. This disorder, as currently understood, can manifest in one of three ways: Individuals with this disorder may be primarily inattentive, may be primarily impulsive/hyperactive, or may present with a combination of both inattention and impulsive/hyperactive behaviors. Research indicates that boys are 3 times more likely than girls to be diagnosed with this disorder and tend to demonstrate the more visible externalizing aspects of the condition, which include overactivity and behavioral dyscontrol. The disorder, which has gained increasing recognition and research over the past 3 decades, affects the educational, social, and behavioral functioning in lives of a substantial number of school-age children, a fact that has significant implications for educational programming in schools.

Research focused on the causes of ADHD has implicated genetic and neurochemical factors as playing a large part in the etiology of this disorder. It appears to be the case that, although environmental factors may play some role in how ADHD symptoms are expressed, maintained, or exacerbated, brain-based and heredity explanations appear to be more likely factors in the expression of ADHD. Although the disorder may be due to these genetic and neurobiological factors, the way in which ADHD is perceived by others is influenced by the context in which the individual is viewed. Children might tend to be viewed as less hyperactive, for example, in a physical education class in which they are required to run and be active as compared with a classroom in which sitting still and focusing on a lesson are required. Therefore, the severity of the disorder can be influenced by the lens of the observer in some cases. For those with a severe form of ADHD, this hyperactivity appears to manifest more clearly across multiple settings.

Some researchers have argued that ADHD is a misnomer for a set of behavioral symptoms resulting from problems with brain-based executive functioning. Executive functioning serves a regulatory purpose for behavior and is thought of as involving higher-order cognitive skills that allow individuals to plan, organize, direct, and control behavior. This capacity also allows individuals to select and persist in adaptive behaviors. The kinds of behaviors that individuals with ADHD have difficulty with are also the kinds of behaviors regulated by this executive system.

In addition to this overseeing function, the executive system assists individuals with time management, memory, and metacognition, or the capacity to think about one's own thought processes and strategies. It is well known that these executive skills develop over the course of our lives and assist in tasks of daily living as well as impacting upon the educational experience of the individual. In fact, research has found that individuals with ADHD may be susceptible to a kind of time blindness thought to be related to deficits in their executive functioning. Self-regulation and self-control, as assisted by skills involving response inhibition, self-regulation of affect, task initiation, and adaptability, are also often problematic for those with ADHD.

ADHD has been recognized as a distinct disorder for many years and has been known by a number of names, including minimal brain dysfunction, hyperactive child syndrome, hyperkinetic reaction of childhood, and attention deficit disorder. The current label of attention deficit hyperactivity disorder refers to one of three variants, with subclassification as primarily inattentive, primarily impulsive, or combined type. The diagnostic criteria for ADHD are outlined in the *DSM–IV*, the primary diagnostic handbook and taxonomy of mental disorders. The diagnosis requires six or more symptoms of specifically outlined symptoms of inattention, six or more symptoms of hyperactivity or impulsivity, or twelve or more symptoms for the combined type. In addition to the manifestation of these symptoms across multiple settings, the diagnosis also requires that the symptoms have been present for at least 6 months with an age of onset before the age of 7. The intensity of the disorder can be described as mild, moderate, or severe. Typically, ADHD is first diagnosed when the demands of the environment make the symptoms most noticeable, which is usually when a child reaches school age. The disorder can manifest itself at younger ages, as well as with more severe forms of the disorder likely identified at earlier ages.

ADHD can be diagnosed by a number of different types of professionals, including psychologists, developmental pediatricians, psychiatrists, and clinical social workers. Each professional will bring different types of understanding to the diagnosis and treatment of the disorder; assessment tools that will assist in the

formulation of a diagnosis also vary according to the clinician's expertise. For example, a pediatrician may feel comfortable with making recommendations for medical and for psychopharmaceutical interventions but may request a more in-depth evaluation of cognitive ability or social emotional functioning that would include a comparison of the referred child's behavior to other children with and without the disorder. In this case, for example, a referral to a school or clinical psychologist might be made for evaluation that would involve normative comparisons.

Although the symptoms of ADHD may seem very familiar to many individuals in a fast-paced, complex culture, and most people feel inattentive, fidgety, or impulsive at different times, it is not the case that we all have ADHD. The diagnosis of the disorder is made as a function of the degree or severity of the symptoms and the extent to which the symptoms have an impact on the person's daily functioning. So, while the nature of the disorder is such that we may all have *some* of the characteristics of ADHD, the actual diagnosis is made based on a variety of variables as well as historical and developmental factors that have a significant impact on the individual's functioning and may range from very mild to very severe. The kinds of information required for such a diagnosis of ADHD are likely to involve interview, a review of the individual's medical and psychiatric history, observations, and rating scales. The diagnostic process should involve the collection of relevant data across multiple settings, multiple sources, and/or multiple informants. Broadband and narrowband behavioral rating scales may be used toward this end. Broadband rating scales assess a broad range of behaviors that might be associated with many types of disorders. Narrowband rating scales target, more specifically, the types of behaviors associated with ADHD. Other formal psychological tests may also be useful in pinpointing an individual's strengths and weaknesses.

Because the age of onset of ADHD is before or within the typical school age, many students with ADHD experience academic difficulties. Inattention can interfere with the acquisition of new knowledge in school and with the retention of already learned concepts. Disruptive behaviors can also interfere with learning and create secondary social or interpersonal problems for students. The research indicates that children diagnosed with ADHD are at greater-than-average risk for experiencing a range of academic difficulties, relationship problems, and conduct problems because of their difficulties with sustained attention

and behavioral disinhibition. Boys are more likely to be diagnosed with this disorder because of the disruptive behaviors they demonstrated within the classroom setting. Inattention, the primary symptom of ADHD in girls, is oftentimes less likely to catch the eye of teachers and caregivers than the disruptive or impulse control problems associated with the clinical presentation of boys.

In addition to academic challenges, the problems that individuals with ADHD experience can change over time due to development. As children grow, increasing behavioral competencies or skills emerge that may affect the presentation of symptoms. In a positive view, for example, increasing capacity to reflect on their own thinking (metacognition), which occurs over the years of child development, may assist the impulsive child in developing effective learning strategies. As teenagers get older, they may expect to engage in increasingly complex psychomotor tasks such as driving—an important rite of passage for many teens. However, adolescent and young adults with ADHD appear to have more automobile accidents and problems with speeding, resulting in more traffic violations and suspended licenses than typical peers, creating additional potential problems associated with these changes in development.

Once a diagnosis of ADHD is made, intervention is often needed to manage the behavioral and cognitive symptoms of the disorder. The most common treatments for ADHD include the use of psychostimulant medications, behavioral programming, psychosocial interventions, and parent training. Psychostimulant medications have been used for many years to treat children with behavioral disorders. In fact, the first use of stimulants in the United States was at the Emma Pendleton Bradley Home for Children in Rhode Island in 1937. The stimulants traditionally used to treat ADHD have included medications such as Ritalin and Dexedrine and are the best researched of any medication for childhood behavioral disorders. A more recent development in the psychopharmaceutical treatment of ADHD is the development of a norepinephrine reuptake inhibitor atomoxetine (Strattera), which appears to hold promise in the treatment of ADHD and holds little risk for abuse and other negative side effects.

Behavioral programming is a form of treatment of the symptoms of ADHD that is based on the principles of applied behavioral analysis and operant conditioning. In this form of treatment, a student's behaviors are viewed in the context of antecedents and consequences

that increase or decrease the likelihood that a particular behavior will be repeated. These antecedent, or consequence, conditions are then experimentally altered to determine what conditions are perpetuating the behaviors. A behavioral management plan is then created for implementation with the student either at home or at school. The plan might include altering the environment in specific ways, such as posting rules so that the student is reminded about behavioral expectations, or moving the student to a place in the classroom that is likely to allow for visual contact with the teacher, for example.

Research indicates that when behavioral approaches are intensive and administered with consistency and in the way in which they were designed, significant improvements can be made in some of the symptoms of ADHD. There also appear some promising results with the use of self-monitoring techniques for the management of symptoms of ADHD. Self-monitoring techniques involve the application of operant behavioral principles with the monitoring of progress to be done by the client or student.

A number of studies have been conducted to evaluate the effectiveness of medical versus behavioral treatments of ADHD. A recent, large, and well-regarded study indicates that when comparing medication management and behavioral interventions, either medication management alone or a combined medication/behavioral approach appears to be most effective. In fact, in terms of academic performance, anxiety, oppositionality, parent–child relationships, and social skills, these combination approaches appear to be most effective.

Other research has examined the effects of psychosocial interventions and parent training on the symptoms of children with ADHD. Improving knowledge about the disorder and clarifying myths about treatment and prognosis are major goals in psychosocial interventions. Because students with ADHD might have associated problems with relationships as a result of their impulsivity, teaching social skills and conflict resolution strategies may be an important element in psychosocial training. Training parents in the treatment of ADHD, using communication and problem-solving skills, as well as behavioral management principles, also appears to be a significantly effective intervention as well.

There are a number of myths about ADHD that are worth noting. Some people believe that people with ADHD are not able to "pay attention," as its name suggests. This, however, is one of the reasons for the inadequacy of the diagnostic label. Individuals with ADHD may have difficulties with sustained attention but are not attention-less as the name implies. ADHD is also not a disorder related to what we traditionally think of as intelligence. Many very bright, capable, and productive individuals struggle with this disorder. In terms of treatment myths, some people are concerned about using "drugs" to treat this behavioral syndrome with the very genuine concern that medication will set precedence for future or illegal drug-taking behavior. In fact, research indicates that individuals who are properly treated for ADHD with safe and effective medications are less likely to abuse drugs and alcohol. Although intuitively appealing, diet and nutritional interventions have failed to show a long-term or significant impact on the treatment of ADHD.

In the past, ADHD was assumed to ameliorate into adolescence. We now know that the symptoms of the disorder often persist well into adulthood. Epidemiological studies report rates of persistence of the disorder from 33% to 66%. Adults with ADHD can continue to experience functional impairments due to their symptoms and may continue to require treatment over the course of their adult lives. Although some of the executive skills involved in planning and organization that are deficient in individuals with ADHD will emerge and develop over time, it seems to be the case that these individuals continue to struggle with ADHD-like behaviors that can affect their daily lives and careers. For this reason, identification of the disorder and early intervention or treatment are important in minimizing the long-term impact of the symptoms of this disorder over the life span.

Mary Ellen Tillotson

See also Applied Behavior Analysis; Behavior Disorders; *Diagnostic and Statistical Manual of Mental Disorders*

Further Readings

Barkley, R. A. (2006). *Attention-deficit hyperactivity disorder: A handbook for diagnosis and treatment* (3rd ed.). New York: Guilford Press.

Dawson, P., & Guare, R. (2004). *Executive skills in children and adolescents: A practical guide to assessment and intervention.* New York: Guilford Press.

DuPaul, G. J., & Stone, G. (2003). *ADHD in the schools: Assessment and intervention strategies* (2nd ed.). New York: Guilford Press.

Kopp, C. B. (1982). Antecedents of self-regulation: A developmental perspective. *Developmental Psychologist, 18*, 199–214.

Lyon, G. R., & Krasnegor, N. A. (1996). *Attention, memory and executive function.* Baltimore: Brookes.

Smallwood, D. L. (Ed.). (1997). *Attention disorders in children: Resources for school psychologists.* Bethesda, MD: National Association of School Psychologists.

AUTISM SPECTRUM DISORDERS

Autism, autism spectrum disorders (ASDs), and related pervasive developmental disabilities are neurological disorders that involve primarily problems of communication, socialization, and behavior. These terms are commonly used to describe the same disabilities; hence, the term *ASD* will be used hereafter. Individuals diagnosed with ASD express the disability in a variety of ways, and a wide range of abilities, strengths, and limitations are common. This once very rare condition of childhood is now commonly diagnosed. ASD has been one of the most researched conditions over the past few decades, and much progress has been made in understanding and supporting persons with the disability. The exact causes are not fully understood, although poor parenting has been thoroughly disproven as a cause.

Even when compared with other disabilities, ASD is an enigma. Children and youth identified as having ASD present highly individualized characteristics that set them apart from their typically developing peers and their peers with other types of disabilities. Some individuals with ASD have near- or above-average cognitive and language abilities with evidence of their disability manifested in the form of subtle social peculiarities. Others have significant cognitive impairments, limited or no expressive language, and severe behavioral and social abnormalities. Individuals with ASD sometimes demonstrate isolated abilities and highly developed splinter skills that contribute to the syndrome's mystery. Fierce debates over the causes of ASD, intervention choices, and educational programming have also been prominent in the recent history of the disability.

Nature and History of Autism Spectrum Disorders

One can logically speculate that ASDs have existed for centuries. Old tales and legends of children and adults with characteristics similar to today's ASDs are scattered throughout history. Stories of the "Holy Fools" of Russia, odd monks, mysterious-acting children, and other individuals who possibly had autism or a related condition can be found in several textbooks. The amount of historical information on ASD is limited for a number of reasons: Families may have kept individuals with ASD from contact with the public; those with mild disabilities may have been perceived to be peculiar, albeit not disabled; some individuals who exhibited bizarre behavior might have been considered possessed or under the influence of magic or spells; some persons with disabilities might have been left to perish in the wild or killed; and medical conditions associated with ASD (e.g., seizures) would have likely resulted in the early death for some.

The words *autism* and *autistic* come from the root Greek word *autos* meaning "self." In this connection the word *autism* was probably first used in the early 1900s by Swiss psychologist Eugene Blueler to describe children who had schizophrenic symptomology and had difficulties relating to others. Major recognition for identifying modern-day notions of ASD is credited to Leo Kanner. In his 1943 seminal work, Kanner described a unique group of children whose behavioral anomalies made them qualitatively different from other children with identified disabilities; he used the term *autistic* to describe them. According to Kanner, these children manifested similar abnormalities from infancy or early childhood, including (a) an inability to relate normally to other people and situations; (b) delayed speech and language development, failure to use developed language for communication purposes, and/or other speech and language irregularities such as echolalia, pronoun reversal and misusage, and extreme literalness; (c) normal physical growth and development; (d) an obsessive insistence on environmental sameness; (e) an extreme fascination and preoccupation with objects; and (f) stereotypic, repetitive, and other self-stimulatory responses. The characteristics of autism as first described by Kanner over half a century ago have been revised, refined, and broadened in recent years, yet current definitions and conceptualizations of ASD continue to reflect many of his original observations.

Working in the same era as Kanner, Viennese medical student Hans Asperger, in 1944, worked with a group of boys whom he described as having an "autistic psychopathology/personality." The children Asperger worked with not only had similarities to Kanner's children but also manifested major differences, such as not having clinically significant

cognitive delay. Because of the language barrier and World War II, Asperger's work was not translated into English until 1981.

Although Kanner referred to genetics and biology early on in his work, he and others soon speculated that the cause of autism was poor parenting and breakdowns in the emotional bonds between parent and child. Bruno Bettelheim, working at the University of Chicago, spoke and wrote about the "refrigerator mother" and emotional difficulties of the parents that led to children having autism.

Perspectives on autism would change in the early 1960s. Bernard Rimland, a researcher and parent of a child with autism, challenged the notion that autism was caused by poor parenting and instead advanced that the disability's etiology was biological and genetic. By the mid-1970s, the acceptance of a psychogenic cause for autism was being discounted, and the neurological nature of the disability was widely accepted.

Andreas Rett made a discovery in 1966 that would eventually add a condition to the autism classification. Rett observed several girls in his practice who had characteristics similar to autism, albeit with significant differences. This new disability would ultimately be named after him—Rett's disorder.

In 1980 autism was added to the American Psychiatric Association's *Diagnostic and Statistical Manual of Mental Disorders* (Third Edition) *(DSM–III)*, replacing the category of early infantile autism. By 1987 another diagnostic label, pervasive developmental disorder–not otherwise specified (PDD–NOS) was added to the *DSM–III*. PDD–NOS was added to the continuum as a means of diagnosing individuals who had some characteristics or more mild traits of autism but not enough symptomology for a full diagnosis of autism. In the mid-1990s childhood disintegrative disorder, Rett's disorder, and Asperger's disorder were added to the *DSM* (1994). Autism was now officially described as a spectrum of similar neurological disorders.

ASD has been one of the most researched conditions over the past three decades, and much progress has been made in understanding and educating students with the disorder. Although poor parenting has been thoroughly disproven as a cause of autism, the exact causes are not fully understood. There are many theories to explain autism, including biological, psychological, cognitive, and affective. Some models are purely genetic and biological in nature, whereas others include environmental causation or a combination of factors.

Definitions and Conceptualizations of Autism and Autism Spectrum Disorders

Commonly used definitions of ASD include those of the American Psychiatric Association (i.e., *DSM*) and the definition advanced by the Autism Society of America. These and other definitions of autism are discussed.

American Psychiatric Association

The most widely used definition of autism in the United States is that advanced in the most recent version of the *DSM* (*DSM–IV–TR*), published in 2000 by the American Psychiatric Association. The *DSM* is an extensively used clinical practice guide; it classifies autism as a pervasive developmental disorder (PDD). In this context, children and youth identified as having a PDD manifest severe and pervasive disabilities in several areas of development, including social interaction reciprocal skills, communication, or narrowly defined and unusual and stereotyped behavior and interests. Such behavioral patterns are demonstrated in the first few years of life and are clearly abnormal relative to a given child's mental age or developmental level.

Subcategories of PDD include autistic disorder, childhood disintegrative disorder, Rett's disorder, Asperger's disorder, and pervasive developmental disorder–not otherwise specified (PDD–NOS).

Autistic Disorder—(Autism)

Autistic disorder, per *DSM–IV–TR* guidelines, is reserved for individuals who display social interaction impairments, communication impairments, and repetitive, stereotypic, and restricted interests and activities prior to 36 months of age. In the majority of cases, children diagnosed as having autism have some degree of cognitive impairment.

Childhood Disintegrative Disorder

In accordance with *DSM–IV–TR* diagnostic standards, children identified as having childhood disintegrative disorder (CDD) typically have behavior patterns similar to those of children with autism. That is, they display the same qualitative social interaction, communication, and behavior/interests impairments as children with autism. However, they differ from children with autism in age of onset of the disability. Children diagnosed with autistic disorder must display

symptoms of PDD before they reach 3 years of age. In contrast, children with CDD experience a period of normal growth and development prior to manifesting social interaction, communication, and behavioral impairments. Thus, following at least 2 years of apparently normal development (but before 10 years of age), children diagnosed as having CDD display a significant loss of skills they previously had acquired. Skill loss falls in at least two of the following fields: language (either expressive or receptive), social skills and social interaction ability, adaptive behavior, bowel or bladder control, play skills and interests, or motor skills. Prognosis for children identified as having CDD is guarded, and typically the social, language, and behavioral challenges associated with childhood disintegrative disorder remain throughout the life span.

Rett's Disorder

Rett's disorder, also called Rett's syndrome, was added to the PDD category of the *DSM–IV* in 1994. In 1999, the gene (MeCP2) that causes 80% of the cases of Rett's disorder was identified. Although Rett's disorder almost exclusively occurs in females, there have been recorded cases of males with this disability. With onset typically occurring by age 1 to 2 years, those diagnosed with Rett's disorder usually develop in an apparently typical fashion during the first five months of life. Onset of the disability is characterized by head growth deceleration; loss of previously acquired motor skills, including purposeful hand movements; stereotypic hand wringing or hand washing; various motor impairments; and social and communication impairments. Loss of these skills is typically progressive and permanent, and prognosis is poor.

Asperger's Disorder

In accordance with *DSM–IV–TR* diagnostic criteria, the essential feature of Asperger's disorder is impaired social interaction. As described earlier, Hans Asperger identified a group of higher-functioning children with autistic-like symptoms in 1944, but this work was largely ignored in the United States until the 1980s. Subsequent to a translation of Asperger's work into English, the syndrome began receiving significantly more attention. This renewed interest has been stimulated, at least in part, by the expansion of the conceptualization of ASD to include individuals with autism-type symptoms who are able to function at a relatively high level. Researchers and practitioners have observed that children with Asperger's disorder are usually able to speak fluently by the time they enter school, albeit their initial language development is sometimes reported to be slow. In spite of their speaking fluency, individuals with Asperger's disorder are noted for having an odd communication style. Many individuals with Asperger's disorder are also reported to be very interested in other people. However, in spite of this interest, they tend to be socially awkward and socially unskilled throughout their lives. Although the exact prevalence of Asperger's disorder is unknown, it appears to be a relatively common form of ASD.

As with other subtypes of ASD, significant elements of the *DSM* diagnostic criteria for Asperger's disorder fall in the areas of *social interaction impairment* (e.g., eye contact difficulties, difficulty in recognizing and reading facial expressions and body language, inability to develop and maintain peer relationships, difficulty in relating and interacting with others) and *stereotypical and restricted patterns of interest and behavior* (e.g., limited, stereotypical, and abnormal interest patterns and aberrant and stereotypical movements, such as flapping). *DSM* diagnostic criteria for Asperger's disorder also include an absence of language, self-help and cognitive delays, a lack of environmental curiosity, and significant occupational and/or social impairment.

Pervasive Developmental Disorder–Not Otherwise Specified

The fifth subtype of pervasive developmental disorder identified in the *DSM–IV–TR* is pervasive developmental disorder–not otherwise specified (PDD–NOS). This somewhat vaguely defined terminology refers to children who evidence significant and pervasive problems in developing reciprocal social interaction skills, verbal or nonverbal communication skill deficits and problems, and/or stereotyped behaviors and interests. The diagnosis of PDD–NOS is used when the criteria for other forms of ASD or other disabilities are not met.

Autism Society of America Definition of Autism

The Autism Society of America (2006) has a widely used definition of autism that is closely aligned with both the criteria used in the *DSM* and Kanner's original work. It also reflects the modern view of autism as a spectrum of neurological disorders. The Autism

Society of America conceptualization includes elements that identify autism as a neurological disorder that affects an individual's capacity to interact and communicate with others. It also identifies autism as a spectrum of disorders that often occur along with other disabilities and that affect individuals differently and to varying degrees of severity.

Other definitions of autism and other related pervasive developmental disabilities include those given by the World Health Organization, the U.S. Department of Education, and advocacy groups such as the Autism National Committee. All of these emphasize the importance of social and language deficits and behavioral anomalies, as well as the acquiring of these characteristics before age 3.

Additional Information and Issues Related to Autism Spectrum Disorders

Prevalence

Estimating the number of children and youth with ASD is difficult. This challenge is due, in part, to difficulties inherent in diagnosing individuals with ASD. There is currently no single reliable and valid test to confirm ASD; rather, diagnosis is made through use of behavioral observations along with tests, interviews, and related methods. Because these methods often rely on subjective interpretations of information and data and clinical judgment, there are differences of opinion among diagnostic professionals in their judgments of individuals who may be considered for a diagnosis of ASD.

Prevalence rates of ASD have changed considerably over the past 30 years. In 1966 it was estimated that autism occurred approximately 4 to 5 times per 10,000 births, and this figure was used through the 1980s. Current conservative estimates are now at about 30 per 10,000; more liberal projections are at about 60 per 10,000. The Centers for Disease Control and Prevention in 2006 noted ASD prevalence at 1 in 166.

Evaluations of school population data of students receiving services under the ASD label also reveal an increasing trend. From the 1991–1992 school year to the 1999–2000 school year, the U.S. Department of Education estimated that there was a 1,108% increase in those receiving services under the autism label in the United States. The exact prevalence of ASD is unknown, yet the fact is that there are significant numbers of children being diagnosed today, and therefore,

appropriate early intervention and continuing services, support, and care are of obvious critical need.

Gender Issues in Prevalence

ASD has consistently been found to be more common among males than females. Most long-standing gender comparisons have suggested that males with autism outnumber females at a ratio of 3:1 to 5:1. Asperger's disorder is also significantly more common among males. Some researchers have even posited that autism may be an extreme variant of the "systemizing" male brain, and thus this would be explanatory of the unequal gender ratio.

Explaining the Prevalence Increases

For both identified and unknown reasons, prevalence rates for autism are now much higher than in previous decades. Numerous explanations have been offered for this increase. First, the conceptualization of autism as a spectrum of disorders as opposed to just one condition has expanded the numbers of individuals diagnosed with ASD. Since the 1980s there have been four additions to the *DSM*'s PDD umbrella category. Most importantly was the addition of those with ASD who are higher functioning and who are rather different from traditional clinical and historical descriptions of persons with autism. Second, over the past 20 years there has been an enormous increase in research and popular press on ASD. As more is known about the condition, increasingly more parents and professionals notice the differences in development and behavior of children and bring these observations to those who can then diagnose ASD-related disorders. Furthermore, the Individuals with Disabilities Education Act in 1991 created a new diagnostic category for students to receive services under the label of autism. Children who had been receiving services under other categorical labels are now receiving special education and related services under the autism label. Finally, there has been much speculation about several other possible reasons for the increase in autism rates. This includes environmental exposures and toxins such as mercury. Some researchers, practitioners, parents, and advocacy groups have claimed that exposure to various toxins, such as mercury (in the preservative thimerosal), lead, aluminum, and other metals, could inhibit neural pathways and thus possibly increase the risks of developing disorders such as

autism. In February 2005, a resolution was introduced in the U.S. House of Representatives that promoted mercury-free childhood vaccinations. However, as of early 2005, many medical professionals and the Centers for Disease Control and Prevention were advising parents to continue vaccinations of children, and little sound, empirical scientific evidence had been presented to challenge these recommendations. The reasons for the significant increase in ASD are not completely clear at this time. However, it appears that the increase can be attributed, at least in part, to an increased awareness of autism among parents and professionals and the broadening of the *DSM* conceptualization of PDD.

Predisposing Factors

Autism and ASD have been shown to be associated with a number of conditions, such as maternal rubella (particularly when deafness or blindness is present), phenylketonuria, encephalitis, meningitis, fragile X syndrome, and tuberous sclerosis. Seizures are also common among some groups of individuals with autism. Depending on the strictness of definitions and the system used to classify, associated medical conditions have been estimated to co-occur with autism at a rate of 10% to 37%. As noted earlier, previous assertions that certain familial and parental interpersonal factors influence the development of ASD have been disproved.

Genetics

Although parenting styles or abilities do not cause ASD, research does support a genetic component. Studies of families with those with ASD have shown the following: (a) Much higher rates of autism occur in identical twins compared with fraternal twins; (b) a family member of a person with an ASD has a higher chance of developing an ASD than does the typical population; and (c) some autism disorders can be directly linked to a genetically linked condition, such as fragile X syndrome or tuberous sclerosis.

Medical Factors Associated With Autism Spectrum Disorders

Children identified as having ASD vary widely in terms of their health and medical characteristics. In 1943 Leo Kanner described children with autism as exceptionally attractive and healthy, and indeed, a number of individuals with autism fit this description. However, there are also children with ASD who have various medical conditions. One such common condition is seizure disorder, which often develops during adolescence. This problem appears to be most prevalent among person with measured intelligence (IQ) below 50 and among children and youth with a diagnosis of autism, CDD, or Rett's disorder. Children with greater cognitive ability and those with a diagnosis of PDD–NOS or Asperger's disorder have a much less chance of having a seizure disorder. The use of medication is widely used to ameliorate seizures as well as behavioral symptoms associated with autism, including self-injury, self-stimulatory behaviors, aggression, and attention problems.

Prognosis

In 1980, the American Psychiatric Association estimated that approximately 1 child in 6 identified as having autism could be expected to achieve marginal social adjustment, be engaged in competitive employment, and could live independently as an adult. The same group estimated that another 1 in 6 could be expected to make only minimal adjustment, and two thirds could be expected to remain severely impaired and unable to live independently. More recent assessments of prognosis suggest that early intervention in combination with a coordinated educational program bodes well for the long-term outlook for individuals with ASD.

The long-term prognosis for individuals with ASD is difficult to determine because of the diverse symptoms related to the disorder and the variance in each individual's abilities as well as their external support systems. As noted previously, even though the ASD condition is considered to be a lifelong disability, some children and youth with ASD become independent adults who show only minimal signs of the major characteristics of their disability. However, even among higher-functioning persons, the social awkwardness or difficulties usually associated with ASD will likely persist. Generally, intelligence (IQ) and language skills are most directly associated with long-term prognosis. It is important to note that these outcome predictors are more reliable when used to assess the outcome of persons with ASD as a group rather than as individuals.

Interventions and Treatments for Children and Youth With Autism Spectrum Disorders

Fierce debates over intervention and treatment choices and educational programming features have been consistent elements in the recent history of ASD. Many parents and professionals perceive ASD to be such a unique disability that they recommend that teachers use specialized and exclusively ASD-oriented intervention, methods, curricula, and programs. Books, magazines, newspapers, and television programs routinely have reports that promote the ever-increasing education and intervention choices for students with ASD.

Professionals and parents have had a particularly difficult time identifying and using the most effective methods from seemingly endless intervention and treatment choices. There is strident debate over which methods bode best for individuals with ASD, and there is often a paucity of scientific evidence to guide choices in this important area. The ASD field is well known for using, and considering use of, interventions and treatments that lack scientific support. A number of professionals and parents have pointed out that the allure of interventions and treatments that lack scientific support is understandable. That is, these methods frequently promise hope for positively responding to a lifelong disability that lacks not only a clear etiology but also a clearly effective treatment plan. Accordingly, it is understandable that professionals and parents who are given opportunities to use methods and treatments that promise dramatic improvements, even if the approach being considered lacks scientific validation, may be willing to "take a chance" and consider using techniques and strategies that lack scientific validation. Unfortunately, these methods have all too frequently been found to be ineffective.

In spite of lack of agreement on specific effective practices, there is nevertheless widespread general recognition of the need for consistent and appropriate use of scientifically supported methods. There is also a growing body of evidence on which intervention and treatment methods have the best records. These strategies and tactics are not universally used, and they can be expected to confer significant benefit only when properly tailored to fit individual student needs and when applied consistently and systematically with fidelity by well-trained and knowledgeable personnel. Currently there is strong agreement among professionals and parents that identifying and using effective practices with children and youth diagnosed with ASD are paramount. Thus, the current issue is not one that revolves around the relative importance of effective practice use but rather on identifying those methods and techniques that have effective qualities and properties.

The task of choosing and using the most effective treatment and intervention methods with children and youth with ASD is frequently complex and challenging. There does not appear to be a single best-suited and universally effective method or program for all learners with ASD. However, it is clear that there are effective methods that can be used to form the foundation of programs for students with ASD and that there are methods and strategies that are generally associated with desired outcomes. The best programs for students with ASD appear to be those that integrate a variety of objectively verified practices and that are designed to address and support the needs of individual children and youth with ASD.

Richard L. Simpson and Paul G. LaCava

See also *Diagnostic and Statistical Manual of Mental Disorders*; Individualized Education Program; Individuals with Disabilities Education Act; Special Education

Further Readings

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., Text rev.). Washington, DC: Author.

Autism Society of America. (2006). *Autism source.* Available from http://www.autismsociety.org

Committee on Educational Interventions for Children with Autism, Division of Behavioral and Social Sciences and Education, National Research Council. (2001). *Educating children with autism.* Washington, DC: National Academy Press.

Frith, U. (Ed.). (1991). *Autism and Asperger syndrome.* Cambridge, UK: Cambridge University Press.

Klin, A., Volkmar, F., & Sparrow, S. (2000). *Asperger syndrome.* New York: Guilford Press.

Koegel, R., & Koegel, L. (2006). *Pivotal response treatments for autism: Communication, social and academic development.* Baltimore: Brookes.

Maurice, C., Green, G., & Luce, S. (1996). *Behavioral intervention for young children with autism: A manual for parents and professionals.* Austin, TX: PRO-ED.

Simpson, R. L., deBoer, S., Griswold, D., Myles, B., Byrd, S., Ganz, J., et al. (2005). *Autism spectrum disorders: Interventions and treatments for children and youth.* Thousand Oaks, CA: Corwin Press.

Wallis, C. (2006, May 15). Inside the autistic mind. *Time*, pp. 43–51.

Zager, D. (2005). *Autism spectrum disorders: Identification, education and treatment.* Mahwah, NJ: Lawrence Erlbaum.

Aversive Stimuli

Aversive stimuli can be defined as any cues or events that produce negative emotional feelings or a negative outcome. Any stimulus can potentially be considered aversive, because it is the production of the undesirable feeling associated with the stimulus that determines whether it is aversive or not. In general terms, the aversive stimulus most often attenuates or eliminates behavior that is paired with that aversive stimulus. As such, aversive stimuli are importantly related to various principles of learning and have major implications for a number of basic and applied research and therapeutic settings.

Aversive stimuli are a key element to escape learning (i.e., the type of learning that is based on negative reinforcement). For instance, animals can learn to escape electric footshock by pressing a bar or moving to a different part of their environment. Aversive experiences are also very important for human learning. The human nervous system has evolved so that even newborns are prepared to find certain types of sensory experiences aversive, so that those aversive stimuli will be avoided thereafter. A wealth of research data indicate that any number of additional stimuli that precede the aversive stimulus can, themselves, become conditioned aversive stimuli. The presentation of the conditioned aversive stimulus can cause a number of behavioral responses that reflect avoidance of, or escape from, the conditioned stimulus. A response that prevents the occurrence of an aversive stimulus, because of the presence of a conditioned aversive stimulus, is called *avoidance conditioning.* Avoidance learning generally occurs after escape learning has occurred, such that the response is eventually made before the aversive stimulus is encountered, so that the aversive stimulus will be entirely avoided. The relationship between escape and avoidance learning is highlighted in the two-stage theory of learning: The first stage involves classical conditioning of fear, and the second stage involves operant conditioning and the reduction of fear by avoidance responding. The negative emotional reaction to aversive stimuli is caused by the activation of parts of the limbic system of the brain, including the amygdala, hypothalamus, and anterior cingulate cortex.

In many instances, an aversive stimulus is the same as a punishing stimulus. Indeed, the use of an aversive stimulus is, in many instances, synonymous with punishment. Unlike positive and negative reinforcement, which act to increase the likelihood of a response, punishment acts to reduce the likelihood of a response, or behavior, occurring again. In other words, punishment produces the opposite outcome to reinforcement. In both experiments and applied settings, punishment may well suppress unwanted behavior—but usually only temporarily and only in situations or circumstances in which the punishment seems unavoidable. Punishment can also lead to additional behavioral issues, such as increased aggressiveness or other exaggerated emotional behavior.

Aversive stimuli potentiate certain types of learning in virtually all animals, as demonstrated by the unique ability of noxious (aversive) sensations to cause long-lasting and often permanent avoidance of the related stimuli. The Garcia Effect is the avoidance of a novel food item that has been associated with illness after only a single exposure. This type of one-trial learning enables the organism to learn the consequences of consuming that food and thus avoid poisonous food sources in the future. In recent years researchers have probed the central nervous system mechanisms that produce the aversiveness of noxious stimuli. Use of the escape/avoidance methodology uniquely facilitates an assessment of the aversive nature of noxious sensory events. Damage to limbic system structures, such as the anterior cingulate cortex, has thus been shown behaviorally to decrease the noxiousness of aversive stimuli.

Aversive stimuli are used clinically in a type of behavioral therapy called *aversion therapy.* This approach is based on classical (or Pavlovian) conditioning, so that the undesirable behaviors are paired with stimuli that produce undesirable feelings or outcomes. The goal is to attenuate or eliminate undesirable behavior. For instance, in the treatment of alcoholism, the undesirable behavior of alcohol consumption can be paired with a drug called Antabuse, also known by the generic name disulfiram. Disulfiram in the presence of alcohol produces a number of aversive physical sensations, including flushing, sweating, soreness in the head and neck, nausea, and vomiting. Theoretically, the repeated pairing of alcohol consumption with these aversive sensations will decrease alcohol consumption. However, the use of this and other

types of pharmacological aversive stimuli that cause uncomfortable consequences are typically associated with poor compliance. It should be noted that behavioral modification techniques utilizing aversive stimuli are often misunderstood. For instance, many times the same aversive stimulus is associated with either negative reinforcement or punishment. In negative reinforcement, the aversive stimulus precedes the undesirable behavior, whereas in punishment, the aversive stimulus follows the undesirable behavior. In combination with adequate cognitive-behavioral therapy, the use of aversive stimuli to treat substance abuse and other disorders may have utility, but there remains controversy as to the ultimate effectiveness of such an approach.

Perry N. Fuchs and John E. McKenna

See also Behavior Modification; Classical Conditioning; Learning; Operant Conditioning; Stimulus Control

Further Readings

LaBuda, C. J., & Fuchs, P. N. (2000). A behavioral test paradigm to measure the aversive quality of inflammatory and neuropathic pain in rats. *Experimental Neurology*, *163*, 490–494.

LaGraize, S. C., Borzan, J., Peng, Y. B., & Fuchs, P. N. (2006). Selective regulation of pain affect following activation of the opioid anterior cingulate cortex system. *Experimental Neurology*, *197*, 22–30.

Mowrer, O. H. (1960). *Learning theory and behavior.* New York: Wiley.

Skinner, B. F. (1938). *The behavior of organisms: An experimental analysis.* New York: Appleton-Century-Crofts.

B

Always behave like a duck—keep calm and unruffled on the surface but paddle like the devil underneath.

—Jacob Braude

BEHAVIOR DISORDERS

Behavior disorder is a general term used to describe a consistent pattern of abnormal conduct impairing a child's ability to function effectively in one or more facets of his or her life. Abnormal conduct may include typical behaviors observed at a developmentally atypical rate or behaviors that are relatively uncommon or bizarre. Careful, multimodal, multi-informant evaluation by a professional is warranted to determine the presence of a behavior disorder. In this determination, a child's conduct is evaluated as to whether or not it deviates from the norm.

One standard that helps establish what is considered normal is whether or not the behavior is developmentally appropriate (i.e., typical for a child of that particular age). For instance, it would be developmentally within limits for a 6-year-old child to have difficulty staying in his seat if a teacher chose to lecture for 20 minutes. A 16-year-old, however, generally should be able to perform this task with minimal difficulty.

Another standard used to characterize behavior as disordered is whether or not the behavior defies social or cultural norms. For example, a child is expected to run around at recess or in gym class. Conversely, it is not acceptable for a child to run around during math class. Some behaviors may be considered inappropriate regardless of context, such as yelling at a teacher or principal. In establishing what is considered normal, professionals should be aware and respectful of cultural influences on behavior. For instance, a young man repeatedly refusing teacher's requests to remove his hat during class may be considered to be defiant. However, if the young man is Jewish and is refusing to remove a yarmulke, his behavior becomes understandable because the teacher's request would conflict with the individual's cultural values.

Aside from being considered aberrant, behavior disorders must also cause the child distress or impairment in at least one area of functioning, such as school, home, or social relationships. For instance, it is atypical for a young child to be able to play *Beethoven's Fifth Symphony* on the piano. Despite this behavioral abnormality, this child would certainly not be considered as having a behavior disorder, because his tendency to execute this action does not affect his ability to function at school or at home. Second, assuming the child is not continuously compelled to play the piece, it probably will not cause him much distress. It is especially important to identify behavior disorders when impaired functioning places a child at risk for harm or loss of freedom (i.e., incarceration). Examples of this are a child who bangs his head when frustrated or an adolescent who compulsively steals. Early identification of these behaviors allows professionals the time to implement steps necessary to better ensure the child's safety and maintenance in the community.

A variety of methods are available to identify behavior disorders. Typically, an adult, such as a parent or teacher, brings the behavior to the attention of a professional. The professional then uses a variety of measures and the reports of multiple informants to assess the child. These measures should include interviews, questionnaires, and direct observation. The technique of using a number of informants and assessment methods allows the professional to be more accurate in his or her determination by ruling out miscalculations such as those due to reporter biases and environmental causes of behavior. To the extent that the child is further classified, to guide service implementation, depends on the context in which the child is being treated.

Within a therapeutic context, if warranted, a child may receive at least one of a variety of diagnoses from the *Diagnostic and Statistical Manual, Fourth Edition, Text Revision* (*DSM–IV–TR*). Behavior disorder in this circumstance encompasses any of the childhood diagnoses that include a high incidence of deleterious behaviors or psychological symptoms, such as anxiety disorders, attention deficit hyperactivity disorder (ADHD), disruptive behavior disorders (e.g., conduct disorder, oppositional defiant disorder), mood disorders (e.g., dysthymic disorder, major depressive disorder, bipolar disorder), or developmental disorders (e.g., Asperger's disorder, autism). Many medical conditions, such as various neurological disorders, may also have behavioral symptoms that place youth in the behavior disorder category. Additionally, a child may not have a behavior disorder, per se, but may exhibit externalizing behaviors as a result of persistent frustration in experiencing symptoms of another diagnosis, such as a learning disorder.

Within school settings, behavior disorders are determined when a child's behavior is having a negative impact on his or her functioning within this environment. In the instance that these behaviors are so severe that additional services are deemed appropriate, the child may be found eligible for a 504 Plan or, in instances of greater impairment, an Individualized Education Program (IEP). Under current special education legislation, some psychological and medical diagnoses are categorized under labels that match the educational exceptionalities (e.g., autism, traumatic brain injury, mental retardation). Most diagnoses that are considered behavior disorders, however, fall under the special education label *emotional disturbance* (ED). Using the *DSM–IV–TR* terminology mentioned in the previous paragraph, ED typically includes anxiety disorders, disruptive behavior disorders, mood disorders, and some developmental disorders (e.g., Asperger's disorder). Alternatively, ADHD and some lower-incidence diagnoses are classified under the special education label *other health impairment.*

Common outpatient clinic approaches to intervening with behavior disorders include therapy focusing on coping skills or interpersonal problem solving and family therapy addressing communication or providing structure and consequences for the child. Within educational settings, informal and formal (e.g., 504 Plan, IEP) supports include reinforcement for desired behaviors, response-cost behavioral systems, meaningful consequences for disruptive behaviors, small-group social skills instruction and practice, and specialized settings that promote coping skills.

Eric R. Benson and Camille J. Randall

See also Attention Deficit Hyperactivity Disorder; Autism Spectrum Disorders; Behavior Modification; Individualized Education Program

Further Readings

Marsh, D. T., & Fristad, M. A. (Eds.). (2002). *Handbook of serious emotional disturbance in children and adolescents.* New York: Wiley.

Schroeder, C. S., & Gordon, B. N. (Eds.). (2002). *Assessment and treatment of childhood problems: A clinician's guide* (2nd ed.). New York: Guilford Press.

BEHAVIOR MODIFICATION

Behavior modification is a technique used for altering an individual's behavior and is based on the principles of learning theory. Using this technique, a behavior can be modified with respect to its frequency, duration, or intensity. Behavior modification is a comprehensive approach that can be applied to behaviors that occur too often (e.g., behavioral excesses), too infrequently (e.g., behavioral deficits), and with behaviors that are both observable (e.g., overt) and those that are not directly evident (e.g., covert). Interventions using behavior modification have produced positive results when delivered to individuals diagnosed with anxiety disorders, attention deficit hyperactivity disorder (ADHD), behavioral and emotional disabilities, and developmental disabilities.

Foundations of Behavior Modification

Classical and Operant Conditioning

Classical conditioning also referred to as *respondent* or *Pavlovian conditioning*, modifies the occurrence of a behavior by pairing two stimuli together to produce a similar response to both. In the famous example, Ivan Pavlov paired an unconditioned stimulus (e.g., meat powder that cued salivation in a dog) with a neutral stimulus (e.g., a bell, which normally would not stimulate salivation). Eventually, the bell became a conditioned stimulus once the dog learned to salivate solely by the sound of the bell.

The theory of operant conditioning, originally put forth by B. F. Skinner, offers a different method for modifying behavior. In contrast to classical conditioning, which focuses on reflexive or involuntary behaviors, operant conditioning modifies voluntary or learned behaviors. Operant conditioning emphasizes *antecedents* (e.g., events that occur prior to behavior) and *consequences* (e.g., events that occur after the behavior), both of which affect the probability that a certain behavior will occur. Whereas antecedents provide a cue for a behavior to occur, behavior is believed to be governed primarily by the consequence of the behavior. That is, the probability of a behavior taking place in the future is determined by what happens after the behavior has occurred. If the consequence of a behavior is desired by the individual, then he or she is more likely to perform the behavior in the future. If the consequence is undesired, then it is less likely that the behavior will occur again. There are two categories of consequences within operant conditioning, *reinforcement* and *punishment.* Consequences are reinforcing if they increase the likelihood of a behavior's occurrence in the future; alternatively, consequences are punishing when they reduce the probability of future occurrence.

Social Learning Theory

Behavior modification is also founded upon Albert Bandura's social learning theory. According to social learning theory, an individual's behavior may be affected by observing the behavior of others. Children and adults learn by witnessing the consequences of others' behavior. An underlying principle of social learning theory is modeling, which suggests that people are more likely to engage in behaviors they observe other people perform if those individuals receive desired consequences.

Ecological Systems Theory

An ecological perspective, proposed by Uri Bronfenbrenner, contends that behavior is influenced by conditions existing in the individual's immediate environment, surrounding environments, and experiences occurring in the larger social, economic, and cultural contexts. Within this approach, individual capabilities are assessed in relation to environmental demands. Consistent with this framework, behavior modification strategies also explore how experiences within one's environment affect behavior. Not only are antecedents and consequences viewed within the individual's immediate setting (e.g., classroom), but their effects can be apparent from events that occurred in a previous setting (e.g., home). In particular, *setting events* are experiences in previous settings that set the stage for a particular behavior in a later setting. For example, a child who did not eat breakfast at home may find it harder to stay on-task in class during the school day.

Behavior Modification Techniques

There are two primary ways of altering an existing environment to modify behavior: adding or removing stimuli from the environment. Positive consequences add stimuli to the environment after the behavior has occurred; negative consequences remove stimuli after the behavior has occurred. Along with the positive and negative aspects to consequences, there are two distinct categories of consequences: reinforcement and punishment. Accordingly, there are four types of consequences: (1) positive reinforcement, (2) negative reinforcement, (3) positive punishment, and (4) negative punishment. Each of these consequences can be used to modify behavior. To be most effective, reinforcement and punishment should be implemented *contingent* upon a specific behavior.

Increasing a Desired Behavior

Increasing behavior utilizes reinforcing consequences, or consequences that, when applied, increase the likelihood that the behavior will occur again in the future. By definition, reinforcement always increases behavior. Thus, it is technically incorrect to state that reinforcement does not work; a consequence is not reinforcing if it does not increase the behavior. Furthermore, even though the terms are commonly used interchangeably, the terms *reward* and *reinforcement* are

not synonymous. A reward is given after a behavior occurs; however, unlike reinforcement, it does not necessarily have the effect of increasing the behavior.

Behavior modification uses two types of reinforcement to increase behavior: positive and negative reinforcement. *Positive reinforcement* is a stimulus added to the environment after a behavior occurs that increases the probability the behavior will occur again. Positive reinforcement can include any stimulus, but often takes the form of verbal praise, tangible objects, or social attention. For example, if saying "Good job" to a child increases the number of math problems completed, then the verbal praise positively reinforces the completion of problems. A behavioral principle related to positive reinforcement is the *Premack principle.* Often referred to as "Grandma's rule," the Premack principle states that a low probability behavior is more likely to occur if it is paired with a high probability behavior. This is similar to your grandmother telling you that you can eat your dessert (high probability behavior) after you eat your vegetables (low probability behavior).

Additionally, behaviors can be reinforced by removing stimuli from the environment. *Negative reinforcement*, or escape conditioning, increases a behavior by removing something aversive from the situation. An everyday example occurs when an individual takes an aspirin to alleviate pain from a headache. If the individual continues to take aspirin for future headaches, it is considered negative reinforcement because the behavior of taking aspirin increases as the result of the pain being removed.

Decreasing an Unwanted Behavior

Consequences that reduce the likelihood of a behavior occurring are called *punishers.* In reducing unwanted behaviors, punishers can be added or removed from the environment. To be most effective, punishment should be immediate, continuous, and intense.

Positive punishment occurs when an aversive stimulus that reduces behavior is added to the environment. Spanking a child is an example of positive punishment. The act of spanking is added to the situation and results in a decrease in the child's previous behavior.

In contrast, *negative punishment* reduces behavior by removing a pleasant stimulus after the behavior occurs. A common example of a negative punishment is "time-out." During time-out, a child is removed from a desired setting for a short period of time because of having performed an unwanted behavior.

The removal of the desired stimulus (e.g., parent attention) serves as a punishment because it reduces the child's previous unwanted behavior.

In addition to these forms of punishment, behaviors can be reduced through *differential reinforcement.* Through this technique, an unwanted behavior is ignored, and another behavior that is an alternative to or incompatible with the undesired behavior is reinforced. An example of this procedure is to reduce running in school hallways by reinforcing walking in the hallway. Walking and running are incompatible behaviors: one cannot occur in the presence of the other. Thus, by reinforcing walking, the unwanted behavior of running is reduced.

Reinforcement Versus Punishment

Behavior management programs tend to focus primarily on reinforcement instead of punishment, despite research demonstrating the effectiveness of punishment. The use of punishment is ridden with multiple limitations, including that (a) it does not teach appropriate behaviors, (b) it does not eliminate or counter reinforcement for behaviors, (c) it can become reinforcing to those administering the punishment, (e) it may negatively affect the behavior of other individuals, and (f) some forms of punishment may be considered ethically questionable. A common behavior management technique used in educational settings is positive behavior support, which focuses on developing appropriate behaviors in students through the use of positive reinforcement. Punishment is still used in combination with reinforcement strategies, but it is typically a secondary component to the behavior management program.

Setting Up a Behavior Modification Program

There are five major steps in developing and implementing a behavior modification plan. These steps include developing behavioral definitions, conducting a functional assessment, collecting data, implementing the program with fidelity, and evaluating the effectiveness of the program.

Behavior Definitions

The first, and one of the most important, components of setting up an effective behavioral modification plan is defining the behavior in appropriate

terms. The behavior to be altered should be defined in terms that are specific, observable, and measurable. The definition should be specific so that anyone observing the individual can easily identify the target behavior. Next, the definition must be in terms that are observable to the individual recording the behavior. Finally, the behavior must be defined in a way that can be quantified to evaluate the effectiveness of the program.

Functional Assessment

A major principle within behavior modification is that all behavior is purposeful. Thus, a key component to developing a behavior modification program is to perform a functional assessment of the behavior identified to be altered. A functional assessment gathers information and develops hypotheses for the intended purpose of the behavior. Information gathering includes conducting interviews and performing behavioral observations in naturalistic settings, such as classroom, playground, or home environments. Interviews provide a means of obtaining anecdotal information regarding the strengths and needs of a child and a preliminary assessment of the behavior in different settings. Conducting observations serves as a way to empirically assess the frequency, duration, or intensity of the behavior and to provide an indication of the antecedents and consequences of a particular behavior (e.g., A-B-C assessment). After information from interviews and observations are analyzed, functional hypotheses of the behavior are derived. Through this process a more appropriate alternative, or replacement behavior, can be identified that serves the same function as the behavior to be modified. A term commonly associated with functional assessment is *functional analysis.* Functional analysis is a process that follows the procedures of functional assessment; however, it extends the process by empirically testing the functional hypotheses to scientifically determine the purpose of the behavior.

Data Collection

All procedures within behavior modification programs are founded on data-based decision making. Throughout the process, data are collected through observations and recorded using behavior observation charts. Before an intervention is implemented, data are collected to provide an estimate of the behavior

(e.g., baseline data). Data collection continues through intervention and follow-up phases to determine if the intervention was successful. Effective data collection is dependent on a specific, observable, and measurable definition of the behavior. Consistency is critical in the collection of data throughout a behavior management program. If more than one person is performing observations and collecting data, then procedures to ensure interobserver agreement must be followed. It is important to evaluate whether the individual observers were using the same observation procedures and recording the behavior in the same way.

Treatment Fidelity/Integrity

Treatment integrity refers to the degree to which an intervention was implemented as intended. It is critical that all participants in a behavior management program implement the procedures in a complete and consistent fashion. Many interventions do not produce desired results, solely because of an inability to consistently implement all components of the program. Measures should be taken to effectively evaluate the degree of treatment fidelity that has been implemented.

Evaluation

Finally, behavior management programs should establish procedures to evaluate the effectiveness of the intervention. The process of evaluation should not be considered a one-time event but rather a continuous process throughout the program. In this way, program evaluation provides a means to modify the program based on immediate information. It also allows the steps of the behavioral management program to be cycled through again whenever needed. Evaluation procedures often include graphing the data collected during baseline and treatment phases. The effectiveness of the program is determined based on visual inspection of the graphed data and statistical analysis. Follow-up data should be collected to determine if the effects of the program are maintained over time. Comprehensive evaluations also include an assessment of social validity, or the extent that the program enhanced the child's quality of life.

John Warren Eagle

See also Applied Behavior Analysis; Contingency Contracts; Operant Conditioning; Premack Principle

Further Readings

Cooper, J. O., Heron, T. J., & Heward, W. J. (1987).
Applied behavior analysis. Columbus, OH: Merrill.

Maag, J. W. (2004). *Behavior management: From
theoretical implications to practical applications*
(2nd ed.). San Diego, CA: Singular.

Martin, G., & Pear, J. (2006). *Behavior modification:
What it is and how to do it.* Upper Saddle River, NJ:
Prentice Hall.

BELL CURVE

The term *bell curve* has several meanings. One meaning is that of a statistical distribution, termed the *normal distribution*, which was identified by Carl Friedrich Gauss. Its shape is that of a bell. The other meaning comes from a book title.

This latter meaning of the bell curve relates to one of the most stimulating and controversial books published in the past two decades, written by Richard Herrnstein and Charles Murray, and titled *The Bell Curve: Intelligence and Class Structure in American Life.* This volume engendered a national conversation and even widespread debates about controversial issues, such as racial and ethnic group differences in intelligence test scores, the respective roles of nature versus nature in influencing intelligence, and the meaning of intelligence.

The primary thesis of this 845-page tome is that American society is undergoing a process of social stratification that is being determined by IQ. Herrnstein and Murray posited that there is emerging a "cognitive elite," a social group composed of highly intelligent people constituting the highest socioeconomic class, and that less intelligent people are gravitating to the lower socioeconomic classes. According to Herrnstein and Murray, a major problem with such social stratification based on IQ is that there will likely be less upward mobility for many individuals from lower socioeconomic classes, because highly intelligent individuals tend to marry highly intelligent individuals rather than less intelligent individuals.

The Bell Curve thus addressed the relationship of IQ to class structure in American life. It posited that success in life is largely based on inherited differences in cognitive ability among people. It also presented findings and commentary on the relationship between race and intelligence.

Although Herrnstein and Murray reviewed many empirical studies in their book, the most important source of data that they examined and discussed is the National Longitudinal Survey of Youth (NLSY), a longitudinal survey of American youth with a very large sample (originally 12,686 youth, ages 14–22). These youth were graduates of American high schools between 1980 and 1982. Herrnstein and Murray carefully reported on various extensive analyses of this data source in the various chapters of their book, *The Bell Curve.* Throughout the book, they used analyses of the NLSY data source to attack various programs of social assistance; these attacks, in turn, ignited widespread debates.

Herrnstein and Murray grounded their work on a traditional view of intelligence based on certain propositions:

1. There is a general factor of cognitive ability on which individuals differ.

2. IQ tests measure intelligence more accurately than do aptitude and achievement tests.

3. IQ tests measure what most people view to be intelligence.

4. IQ scores are relatively stable over much of human life spans.

5. IQ tests tend not to be biased against economic, ethnic, racial, or social groups.

6. IQ is substantially heritable, with 40% to 80% of the variation in IQ scores being attributable to hereditary factors.

There are numerous conclusions resulting from the extensive analyses in their work. Regarding poverty, Herrnstein and Murray concluded that IQ is a stronger predictor of poverty than is socioeconomic background. Regarding unemployment, they concluded that low IQ is a greater risk factor than either education or socioeconomic background. Regarding children born out of wedlock, they found that low IQ for a woman increases the likelihood that the woman will have an illegitimate child. Regarding welfare dependency, they determined that low IQ is a prominent predictor of the receipt of public welfare funds among women. Regarding parenting, they found that low IQ among White mothers is a strong predictor of low birthweight among children. Regarding crime, they concluded that low IQ is a risk factor for criminality.

Regarding ethnic differences, they concluded that Americans of East Asian background (e.g., Japanese Americans, Chinese Americans) tend to have higher IQ scores than Euro-Americans who, in turn, tend to have higher IQ scores than African Americans and that the gap in IQ scores between Euro-Americans and African Americans has narrowed by about three IQ points. They also determined that the difference in family income between Euro-American families and African American families is markedly reduced when one controls for IQ. Regarding social problems, they determined that individuals with social problems tend to be low IQ individuals. In general, Herrnstein and Murray contended that intelligence affects many social and psychological attributes of individuals with high IQ, having many more salutary effects on high IQ people than on low IQ people.

If higher intelligence is a protective factor against numerous social ills, then what about efforts to raise intelligence? Herrnstein and Murray reviewed research on efforts to improve intelligence and found that initiatives to improve nutrition, formal schooling, and preschool education may have positive effects on intelligence but that the research on such initiatives is thus far inconclusive. However, they strongly support adoption at birth for children from problematic family settings to good family environments as the major way to improve the intelligence of children.

One noteworthy feature of their book is that the vast majority of the data that they analyzed dealt with Euro-American individuals—especially, Euro-American females. As a result, it is improper to view it as a tome primarily concerned with ethnic and racial group differences in IQ. *The Bell Curve* is, however, an exposition of a theory of cognitive stratification of American society.

Herrnstein and Murray predict certain specific trends in American society. First, the cognitive elite will become increasingly isolated. Second, the cognitive elite and the affluent will merge to become the new upper socioeconomic class. Third, the quality of life for the people in the lower half of the IQ distribution will deteriorate.

To address the social problems resulting from the cognitive stratification of society, Herrnstein and Murray contend that certain changes should occur in government. First, the rules that underlie starting and running a business should be simplified, so that more people from all levels of society can make a viable living and find valued places in society. For example,

the enormous federal tax code should be simplified. Second, the criminal justice system should be simplified, clearly indicating the meanings of criminal offenses and their consequences.

The Bell Curve has engendered much subsequent inquiry on intelligence and even strident academic debate.

William M. Bart and Saa Hoon Hong

See also Intelligence and Intellectual Development; Intelligence Quotient (IQ); Social Class and Classism

Further Readings

Feuerstein, R., & Kozulin, A. (1995). The Bell Curve: Getting the facts straight. *Education Leadership*, 52(7), 71–74.

Herrnstein, R. J., & Murray, C. (1994). *The Bell Curve: Intelligence and class structure in American life.* New York: Free Press.

Jacoby, R., & Glauberman, N. (Eds.). (1995). *The Bell Curve debate.* New York: Times Books.

Neisser, U. Boodoo, G., Bouchard, T., Boykin, A., Brody, N., Ceci, S., et al. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, 51(2), 77–101.

Tittle, C. R., & Rotolo, T. (2000). IQ and stratification: An empirical evaluation of Herrnstein and Murray's social change argument. *Social Forces*, 79(1), 1–28.

BILINGUAL EDUCATION

While there are about 200 recognized sovereign nations, there are well over 6,000 languages spoken throughout the world. Because of increased migration, geographical proximity, and/or political conquest and colonization, few countries today can claim monolingualism as the norm. Moreover, globalization has placed English in a unique role in many school systems throughout the world. Bilingual and multilingual education is one form of schooling that has been developed worldwide in response to this linguistic and cultural diversity.

The terms *bilingual students* and *bilingual education* are sometimes confused. Bilingual children know and use two languages to different degrees. Depending on the nature of access to both languages, as well as attitudes toward the languages, bilingual children demonstrate varying proficiency in their two languages; for example, they may speak both languages

but be literate in only one language. Their bilingual skills and the extent they identify culturally with the two languages may develop and vary over time. Bilingual children may or may not attend a bilingual education program.

This entry outlines bilingual and multilingual education for minority (dominated) language and majority (dominant) language speakers, using examples from nations throughout the world. After introducing some basic definitions of key terms, the entry highlights various models that are traditionally distinguished. The third section addresses issues and trends related to the implementation of bilingual education programs.

Definitions

Simply defined, *bilingual education* is instruction that uses two languages as media of instruction. By extension, multilingual education programs aim for proficiency in more than two languages. These programs are implemented in many different forms in countries all over the world and respond to national and local contexts, student needs, and available resources.

In this entry, *native language* will be used to refer to the language in which the child has been raised, although it must be noted that in multilingual environments this can be more than one language and it may not always be the same variety as the (standard) language variety taught in school. The dominant or societal language is the predominant language used for communication in the students' nation (including government, education, media). It generally has a high-status standard variety that is used and taught in schools. A *second language* is a language learned at a later stage than the native language. This often occurs outside the home through school or the media. *Heritage language* is the language used by a particular ethnic group. *Minority* or *dominated languages* are languages used by language groups who are politically and socially placed in a minority situation but may not necessarily be numerically in the minority. In many school districts in the United States, Spanish is a minority language even though Spanish-speaking students may constitute the largest student group. *Majority language* speakers are speakers of the dominant, or societal, language. They are increasingly a numerical minority in urban schools.

The definition of bilingual education as instruction in and through two languages does not consider foreign language classes as a form of bilingual education because, even though some form of bilingual proficiency is reached, the foreign language is only taught as a subject. This entry also excludes out-of-school efforts for native language maintenance through community-based organizations that complement monolingual education in the societal language; although this situation results in de facto bilingual education, both languages are not used within the same instructional approach. Finally, it must be noted that much research has described bilingual education at the elementary grades, though interest in preschool and secondary models is increasing.

Bilingual and Multilingual Models

Many different classifications and descriptions of bilingual and multilingual education programs exist. Three broad criteria distinguish among the most commonly implemented models: *program goals, target population*, and the *distribution of the languages*. In terms of goals, some programs aim for additive bilingualism or multilingualism; that is, the program's goal is to add one or more new languages to the student's native language. In contrast, subtractive models are bilingual approaches where the main goal is to facilitate the learning of the societal language. A second criterion that distinguishes among various models is the student population. Many programs enroll native speakers of the societal language and teach them a second or third language. Other programs exclusively target speakers of dominated languages, including regional and immigrant languages. Few bilingual programs enroll a dual target population of native speakers of the societal and a dominated language. Some bilingual programs purposefully integrate students from different language groups for certain times of the day or week. The third criterion considers the use and distribution of the two or more languages for instructional purposes. The amount of time allocated to teach language and the subjects to be taught in each language is directly related to the goal and target population. The choice of the language(s) of initial literacy instruction is also a key decision in bilingual and multilingual programs.

Table 1 summarizes some of the most common program labels used in the literature. Bilingual and multilingual programs can be programs within schools, or they can be schoolwide. A program label can represent only a general description of a bilingual or multilingual program. The actual implementation of

Table 1   Bilingual and Multilingual Program Models

Program Label	Language Goals	Target Population	Language Use and Distribution
Multilingual education	Multilingualism	Minority, majority	L1, L2, and L3
European schools	Multilingualism	European Union civil servants	L1, L2, and L3 (and optional L4)
Mainstream bilingual education	Bilingualism	Majority, international	L1 and L2
Canadian immersion programs	Bilingualism	Majority	L2 and then L1
Two-way immersion	Bilingualism	Minority, majority	Long-term L1 and L2 use
Maintenance bilingual/ heritage language education	Bilingualism	Minority	L1 and L2
Bilingual education for the deaf	Bilingualism	Minority	L1 and L2
Transitional bilingual education	Proficiency in L2	Minority	L1 for limited amount of time and L2
Integrated transitional bilingual education	Proficiency in L2	Minority	L1 for limited amount of time and L2

Source: Adapted from Brisk, M. E. (2006). *Bilingual education: From compensatory to quality education* (2nd ed.). Mahwah, NJ: Erlbaum; and Baker, C. (2006). *Foundations of bilingual education* (4th ed.). Clevedon, UK: Multilingual Matters.

each model can vary widely depending on the availability of qualified teachers, resources, official policies, and societal and school (district) attitudes.

Bilingualism for Societal Language Speakers

Programs for native speakers of the societal language are typically enrichment programs, that is, optional programs for parents who want their children to develop high levels of bilingual competence. Generally speaking, these programs cater to children from higher socioeconomic backgrounds and are associated with *elite* or *elective bilingualism.* Elite bilingualism is largely uncontroversial and supported as an enrichment possibility for majority language speakers for personal development and access to global economic opportunities. Two bilingual program options that have been developed for societal language speakers are mainstream bilingual education and (Canadian) immersion programs. Multilingual models are found at the school level (the European School model and other schools) as well as the country level.

Mainstream bilingual education often takes place in the students' native language and an international language, such as English. International schools have traditionally served an elite, often mobile population through monolingual instruction. The increased enrollment of non-English-speaking international and local students has led to more bilingual or multilingual schools where English and the societal language are taught. International schools in Brazil use Portuguese Brazilian in the morning and English in the afternoon and often hire teachers from Brazil and North America.

Probably the best-known bilingual enrichment programs for dominant language speakers are *Canadian immersion programs.* (The term *Canadian* is added to distinguish these bilingual programs from English immersion programs for minority students, which aim for second language proficiency rather than bilingualism.) French immersion programs were developed in the late 1960s in response to demands from middle-class, Anglophone parents living in predominantly French-speaking Montreal, Quebec, to provide their children with the opportunity to develop a functional level of bilingualism. Immersion models differ with regard to the introduction of the second language (early versus late immersion) and the amount of time

spent in the second language (full or partial immersion). In early, full immersion programs, the target (second) language, French, is introduced before the student's first language during the first two years of elementary school. By third grade, the student's native language (English) is formally introduced in the curriculum, and both languages are used for equal amounts of time for the rest of the program. In the case of late immersion, the second language is introduced either at the upper elementary (middle immersion) or at the secondary level (late immersion). In Hungary, English is introduced as a medium of instruction at the secondary level, alongside Hungarian. Canadian-style immersion programs can be found as elementary foreign language immersion programs in the United States in Chinese, French, Japanese, or Spanish.

The *European school model* is a multilingual model originally designed for children of parents who worked for the European Coal and Steel Community. There are now 10 European schools in Luxembourg, Belgium, Germany, Italy, the Netherlands, and Great Britain, primarily enrolling the children of civil servants working for the European Union. The schools follow a common model with two main goals: to maintain and develop the students' native language and cultural identity and to promote a European identity. The European school model begins with extensive instruction in the students' native language and a second language initially as a subject area. By second or third grade, the second language use as a medium extends to subjects that are less language-dependent. Academically complex instruction in the second language is gradually increased until by eighth grade, subjects such as science and history are taught in and through the second language. European identity development is supported through common "European hours" during which students from multilingual backgrounds engage in cooperative and hands-on learning of content taught from a European perspective.

In some European countries, multilingual models are implemented wherein a high prestige language is added to a bilingual education program or three languages are used as media of instruction. Catalan immersion programs for native Spanish speakers include Catalan and Spanish, as well as English as a third language. In Switzerland, private schools offer bilingual streams in German and French. English is used to teach subject matter and cultural activities as of seventh grade, and students can add Latin or Italian

in later years. Luxembourg has official trilingual education for all students enrolled in the school system. Mandatory preschool education begins in Luxembourgish, and this language continues as a language of instruction and as a subject throughout elementary school. German is introduced as a subject the first year of primary school and then intensifies as a subject and medium of instruction until sixth grade. Similarly, French is introduced as a subject in second grade and then increases its role in the curriculum to be the exclusive medium of instruction by the age of 15. Foreign language education is introduced at the secondary level as well, including English, Latin, Spanish, Italian, or Greek. A trilingual policy also shapes the education system in India, where students are expected to learn the regional language, Hindi, and English, regardless of their native language.

Bilingualism for Dominated Language Speakers

Dominated or minority language speakers generally do not have a choice in becoming bilingual. For these groups, learning a second language is necessary for survival and access to the society in which they live as a result of forced migration, conquest, or voluntary immigration. This form of bilingualism is often referred to as *folk* or *circumstantial bilingualism* and, unlike elite bilingualism, it is often controversial in educational policy. In recent years, speakers of regional and indigenous languages have received more legal protection through national and international law. Similar linguistic rights do not yet exist for immigrant or ethnic languages. As a result, the position of indigenous and regional minority languages and ethnic or immigrant languages in school varies.

The consistent repression and rejection of the use of indigenous languages in schools as part of the nation-building process has been documented throughout the world. In the United States, the 300 Native American languages that existed upon European contact have been reduced to only a small number of languages still widely spoken by children today. Until recently, European national governments have largely ignored regional language groups with long-standing historical and cultural roots in the area (such as Catalan in Spain, Frisian in the Netherlands, Welsh in Wales) or even actively prohibited the use of these languages in school, as was the case for Basque under General Franco's regime in Spain. As

a result, fluent speakers of these indigenous languages and regional minority languages have been disappearing rapidly. *Heritage* or *bilingual maintenance* programs have been established to revitalize these indigenous and minority languages to increase the number of speakers, expand the domains where the language is used, or both.

The Māori in New Zealand provide full immersion in the minority language in preschool (through Te Kōhanga Reo programs or "language nests") and elementary school (Kura Kaupapa Māori) before introducing English as a medium of instruction. A similar approach was chosen to revitalize the Hawaiian language. In Bolivia, after many years of repression, Quechua and Guaraní are now used as media of instruction. With an emphasis on community involvement, indigenous languages and cultural practices are included in the curriculum in several South American countries, in addition to the official language, Spanish. In the United States, language revitalization efforts are particularly strong for Native American languages. Teresa McCarthy describes the history of Rough Point, Arizona, a Navajo–English bilingual program. Subjects are taught in both English and Navajo through second grade. As of third grade, English becomes a medium of instruction, and Navajo continues to be used as a vehicle for studying Navajo culture and citizenship.

Language maintenance and revitalization efforts in non-English-speaking countries increasingly include a multilingual component through the addition of a third, high-prestige language such as English, to the regional minority language and the societal language. In the Basque region (Spain), the Basque language is used as a medium of instruction from kindergarten through secondary school. Spanish and English are initially included as a subject at the elementary level; English is then also used as a medium of instruction at the secondary level. German or French are offered as optional foreign languages. In the Netherlands, schools are experimenting with the equal distribution of Frisian and Dutch and the introduction of English as a medium of instruction in fifth and sixth grades.

Bilingual programs for the deaf teach content through sign language and the written form of the societal language. Approaches to language (the use of different sign systems) and the curriculum vary greatly from program to program and is often constrained by teachers' lack of proficiency in sign language. Deaf culture and identity are important components of the bilingual program to counter deficit views of individuals who are deaf or hard-of-hearing.

In most nations, *transitional bilingual education* (TBE) is the model most commonly chosen for ethnic and immigrant languages. In a TBE program, the student's native language is used for an initial period of literacy development and content learning to assist the students' transition to literacy and content learning in the societal language. The goal is not to maintain or develop the students' native language but to provide access to the language of school. Students are typically expected to enroll in the program for 2 or 3 years and then exit the program into a standard classroom or instruction in the societal language only (early-exit TBE). In most African postcolonial nations, bilingual education that includes the student's native language is transitional and often experimental in nature. Other TBE programs attempt to provide a more gradual transition (late-exit TBE). Students begin literacy and content instruction in their native language, with English as a second language instruction. In second grade, formal literacy instruction in the second language is started. In third grade, instruction is increasingly balanced between the two languages across subjects; often the native language is used to preview major concepts and activities which are later taught to students in English. Academic English instruction continues to increase by the second half of third grade and students are expected to exit the program at the end of fourth or fifth grade.

Integrated Bilingual Education Models

Few bilingual models address the need for bilingualism or multilingualism for an integrated minority–majority student population. In most cases, the focus is on avoiding the segregation of minority language students while they are becoming familiar with the new language and culture (integrated TBE). The Foyer model in Belgium provides half of the instruction for Turkish-speaking students in preschool and kindergarten separately and half of their instruction integrated with majority language speakers. By third grade, 90% of the instruction is integrated with majority language speakers, and 10% of the instruction is on native language and culture. Integrated TBE models exist in Scandinavia and the United States, where minority language students are separated for specialized instruction in their native language and

integrated with native English speakers for specific subject areas for varying amounts of time.

The most developed integrated bilingual education model is *two-way immersion* (TWI) education. In the United States, TWI is an integrated model of bilingual education according to which native English speakers and native speakers of a minority language are educated together for most or all of the day and receive content and literacy instruction through both English and the minority language. Its goals include academic achievement, bilingualism and biliteracy development, and cross-cultural competence for all students. TWI programs vary by the distribution of the heritage and the societal language. Some programs start with literacy and content instruction in the minority language for 90% or 80% of the time; others distribute the two languages more evenly.

Issues and Trends in Multilingual Education

The existence of various bilingual education models reflects the basic principle that quality education through multilingual approaches can and needs to occur in different ways to respond to international, national, and local goals and needs. Despite these multiple realities, the term *bilingual education* is often used to serve particular political and ideological ends. Supporters of English-only policies, for example, limit the term to instruction exclusively in the native language. Bilingual education advocates sometimes argue that the term should be applied only to programs in which the two languages have equal status and are used equally for teaching. Both definitions are narrow in scope and exclude the realities and multiple approaches to multilingual development that exist.

The two definitions also reflect a dichotomy that has often hindered systematic research that can inform effective program implementation. Much research on bilingual education has attempted to justify either the use of the native language (in addition to the societal language) or solely the societal language. Whereas this debate has resulted in many, highly contested program evaluation studies, it has contributed little to the understanding of the factors that contribute to positive linguistic, sociocultural, and academic outcomes in schools. To guide program implementation, it is necessary to consider the interaction between theoretical principles and the dynamics of language and cultural practices in real schools and classrooms. Studies have found that effective schooling for language minority students comes from additive bilingual programs that are fully integrated in the school, taught by quality teachers, and guided by knowledgeable leadership. These programs have bilingual curricula that reflect high expectations of students, use bilingual teaching methodology and assessment practices, and have strong parent and community involvement.

The perceived problematic nature of bilingual programs for minority language speakers contrasts sharply with the desirability and value of multilingual proficiency for speakers of the societal language. These conflicting positions exist within diverse nations and point to a double standard in educational language planning. Whereas language majority speakers are encouraged to expand their linguistic repertoire beyond their native language, speakers of minority languages have to struggle to maintain their native languages and prevent them from dying out. The double standard results in a dual education system: one that values and actively promotes bilingualism (for the elite) and one that rationalizes the importance of (monolingualism in) the societal language (for language minority students). France promotes foreign (i.e., official European Union) languages for its native French-speaking population but fails to provide bilingual education options that use immigrant languages such as Turkish.

Bilingual and multicultural education is increasingly shaped by globalization and the role of international languages in education, in particular English but also Chinese, Arabic, and Spanish. In Canada, high school students in Ontario view French as a commodity that can provide access to economic opportunities beyond the local community into the international marketplace. The positioning of English and other international languages within and across nations plays an important role in multilingual education. Competition with English or other international languages can undermine the availability of quality instruction in and through indigenous or ethnic languages in non-English-speaking countries. In Taiwan, the government established policies to promote Taiwanese and English, but an analysis of the actual implementation of the policies showed that teacher training and curriculum and resource development occurred for the English language but not for Taiwanese (or other minority languages). English is also taking over functions and domains traditionally reserved for the societal language, particularly in business and scientific discourse, which may cause

language shift or loss over time. Unequal access to multilingual instruction that supports native language instruction and adds not only the societal language but also English may constrain educational opportunity and social mobility.

Bilingual and multilingual education models challenge monolingual norms and practices that are prevalent today. In the United States, the federal No Child Left Behind Act and recently passed English-only laws in California, Arizona, and Massachusetts rejected the value of multilingualism for its language minority population by insisting on English-only instruction and the assessment of achievement through English. The monolingual bias reflected in these laws views multilingualism as an abnormality rather than a global reality and treats multilingual resources as a deficit rather than a resource. By bridging the divide between elite and minority bilingual education practices, schools can build on and extend existing linguistic and cultural resources to meet current and future language, cultural, political, social, and economic demands and develop multilingual competence for all students.

Ester Johanna de Jong

See also American Indians and Alaska Natives; Bilingualism; Cultural Diversity; Diversity; Immigration; Multicultural Education

Further Readings

Azurmendi, M. J., Bachoc, E., Zabaleta, F. (2001). Reversing language shift: The case of Basque. In J. A. Fishman (Ed.), *Can threatened languages be saved?* (pp. 234–259). New York: Multilingual Matters.

Baker, C. (2006). *Foundations of bilingual education* (4th ed.). Clevedon, UK: Multilingual Matters.

Baker, C., & Jones, S. P. (1998). *Encyclopedia of bilingualism and bilingual education.* Clevedon, UK: Multilingual Matters.

Brisk, M. E. (2006). *Bilingual education: From compensatory to quality education* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.

Cenoz, J., & Genesee, F. (1998). (Eds.). *Beyond bilingualism: Multilingualism and multilingual education.* Clevedon, UK: Multilingual Matters.

Cummins, J., & Corson, D. (1997). (Eds.). *Encyclopedia of language and education: Vol. 5. Bilingual education.* Dordrecht, Netherlands: Kluwer.

De Mejia, A. (2002). *Power, prestige and bilingualism. International perspectives on elite bilingual education.* Clevedon, UK: Multilingual Matters.

Extra, G., & Yağmur, K. (2004). (Eds.). *Urban multilingualism in Europe. Immigrant minority languages at home and school.* Clevedon, UK: Multilingual Matters.

Genesee, F. (2004). What do we know about bilingual education for majority language students? In T. K. Bhatia & W. Ritchie (Eds.), *Handbook of bilingualism and multiculturalism* (pp. 547–576). Malden, MA: Blackwell.

Glenn, C. L., & de Jong, E. (1996). *Educating immigrant children: Schools and language minorities in twelve nations.* New York: Garland.

Howard, E. R., Sugarman, J., & Christian, D. (2003*). Trends in two-way immersion education: A review of the research* (Rep. No. 63). Baltimore: Center for Research on the Education of Students Placed At Risk (CRESPAR).

Johnson, R. K., & Swain, M. (1997). (Eds.). *Immersion education: International perspectives.* Cambridge, UK: Cambridge University Press.

May, S. (2001). *Language and minority rights. Ethnicity, nationalism, and the politics of language.* Reading, MA: Pearson Education.

Nettle, D., & Romaine, S. (2000). *Vanishing voices. The extinction of the world's languages.* Oxford, UK: Oxford University Press.

Strong, M. (1995). A review of bilingual-bicultural programs for deaf children in North America. *American Annals of the Deaf*, *122*, 84–94.

BILINGUALISM

Bilingualism is a major fact of life in the world today. One in three of the world's population routinely uses two or more languages for work, family life, and leisure. Many more make irregular use of languages other than their native one(s), including those who have learned an additional language at school and use it occasionally for specific purposes (e.g., travel). It seems likely that this trend will continue given the current processes of globalization and the growing need to communicate across political and linguistic borders.

The term *bilingualism* can be used to describe an individual phenomenon, as in a person's ability to speak two or more languages. It is also possible to talk about bilingualism as a characteristic of a group or community of people, as bilinguals and multilinguals are usually found in groups, communities, or particular regions. In this entry, the discussion of bilingualism focuses on the individual level. The entry commences with a definition of individual bilingualism, followed by a discussion of related dimensions and characteristics. It then provides an overview of the reported advantages of bilingualism and concludes

with a description of the factors associated with the development of bilingualism in infancy and beyond.

Defining Bilingualism

The concept of bilingualism has broadened since the beginning of the 20th century. Although earlier definitions tended to restrict bilingualism to equal mastery of two languages, it is now recognized that a bilingual or multilingual speaker uses different languages for different purposes, in different contexts, and in communicating with different partners and does not necessarily possess the same level or type of proficiency in each language. The broadening of the concept of bilingualism is largely due to the realization that the point at which a speaker/user of two languages becomes bilingual is either arbitrary or impossible to determine.

Despite its omnipresence, defining bilingualism, or a bilingual individual, proves to be more difficult than it appears. The term *bilingual* describes primarily someone who can function in two languages in conversational interaction. It can also include the many people in the world who use, and have varying degrees of proficiency in, three, four, or more languages. In his definition of *bilingualism*, François Grosjean focuses on the daily use of two or more languages and distinguishes bilinguals, who use two or more languages in daily life, from "dormant bilinguals," who retain knowledge of different languages but no longer use them in daily life.

Describing Individual Bilingualism

According to linguist William Mackey, bilingualism is a behavioral pattern of mutually changing linguistic practices that vary in degree, function, alternation, and interference, or cross-linguistic influence. It is through these four related characteristics that individual bilingualism may be best described.

Degree. Because bilingualism is a relative concept, it involves the question of degree. In describing a person's bilingualism, the most obvious thing to determine is how well an individual knows the languages he or she uses. Because bilinguals are rarely equally fluent in all language skills in all of their languages, this would involve separate assessments of comprehension and expression in both the oral and written forms of each language. In this way, ability in two or more languages is best understood on a dimension or continuum, with dominance and development varying across people. For example, some bilinguals actively speak and write in both languages. This is referred to as productive competence. Others have more passive, or receptive, ability in a language, such that they may understand or read it but may not be able to speak or write it very well. Ability in each domain may be relatively advanced in both languages or may just be developing in a second or third language.

Furthermore, it is necessary to determine a bilingual's mastery of speech sounds or graphics, grammar, vocabulary, meanings, and discourse in each skill for each language, as a bilingual's mastery of a skill (i.e., listening, speaking, reading, and writing) may not be the same at all linguistic levels. For example, a bilingual may have an extensive vocabulary but a poor pronunciation, or a native-like pronunciation but underdeveloped grammar. In effect, proficiency in two sets of related variables, skills, and levels needs to be determined. This framework highlights the complexities involved in defining and understanding individual bilingualism, while at the same time it belies the existence of great within-group diversity.

Bilinguals are sometimes expected to be two monolinguals in one person. Grosjean referred to this as the monolingual, or fractional, view of bilinguals. He argued that the bilingual is not the sum of two monolinguals but has a unique linguistic profile that is a complete linguistic entity, an integrated whole. This is because bilinguals do not usually possess the same competence as monolingual speakers in both of their languages because bilinguals use their languages with different people, in different contexts, and for different purposes; furthermore, levels of proficiency in a language may change over time according to changing circumstances and linguistic contexts. Communicative competence in one of a bilingual's two languages may be stronger in some domains than in others. As Grosjean noted, this is natural and to be expected.

Function. Mackey argued that bilingualism is not a phenomenon of language but a characteristic of its use. Use is an important factor in the development and maintenance of bilingualism, as individuals often use their languages for different functions and purposes; these functions, in turn, affect their degree of proficiency in each language. The domains where each language is acquired and used also vary. Functions of bilingualism may be external or internal.

External functions of bilingualism are areas of language contact that vary by duration, frequency, and pressure. Areas of contact include all media through which languages are acquired and used (e.g., in the home, community, school, radio, television, print materials, and other uses of the printed word).

Home language(s) may reflect or differ from those of the community and other language contacts. Bilingual parents may decide to each use a different language with their children, so that one parent uses language A while the other parent uses language B. This practice is often referred to as Grammont's principle, or "one person, one language," and the effects have been studied to test the theory that two languages can be acquired from birth for the same effort as acquiring one. Community languages include those languages spoken in the bilingual's neighborhood, ethnic group, church group, workplace, and social group. The more prevalent language of the greater community often takes the place of the home language as the most important influence on a bilingual child's speech. This presents many challenges for families trying to raise children bilingually in the majority language and a less prevalent one. A person's language contact in school may be with a language taught as a subject (e.g., foreign language class) or with a language used as a medium of instruction (e.g., two-way immersion or dual-language program). The community's support for, and the prevalence of, both languages are important factors in the development and long-term maintenance of each language. Radio, television, movies, music, newspapers, magazines, and the Internet are powerful media in the maintenance of bilingualism. Access to these media may be the key factor in maintaining one of a bilingual's languages, especially if the other language is the only one spoken in the area.

Duration, frequency, and pressure are important variables in the development and maintenance of a language. The amount of time a person has lived in a community that uses, supports, and promotes a language (or languages) will affect that individual's command of the language(s). Languages that are used as mediums of instruction will provide more contact hours than those that are taught as individual subjects. How often a person uses each language is as important as how long he or she has been exposed to and has been using those languages. Finally, language use is affected by the types of pressures associated with it, including economic, occupational, cultural, political, military, historical, religious, and demographic.

Bilingualism is also influenced by a number of internal functions. These include noncommunicative uses, such as counting, dreaming, and note-taking. Some bilinguals may use one of their languages for most or all of their inner expression, whereas others may use different languages for different types of internal expression. Internal functions also include factors that are likely to influence a bilingual's aptitude for using each language. Sex, age, intelligence, memory, language attitude, and motivation are regarded as important factors in the development and description of individual bilingualism.

Alternation. The alternate use of two languages, or codeswitching, within a conversation is a natural phenomenon of bilingualism. Studies in the past 20 years have shown that codeswitching is a highly complex and sophisticated process governed by certain rules and constraints that do not violate the syntax of either language. Codeswitching involves the skilled manipulation of overlapping sections of two (or more) grammars; there is virtually no instance of ungrammatical combination of two languages in codeswitching, regardless of the bilingual ability of the speaker. For example, bilinguals may codeswitch because they are momentarily lost for words in one of their languages, or do not know a word or phrase in one language, or can express an idea more effectively in the other language. Rather than demonstrating confusion or a deficit in their languages, bilinguals codeswitch in practical and purposeful ways, usually when in the company of other bilinguals, in an attempt to express themselves with the vocabulary and grammar available to them and/or with more accurate language. Codeswitching is also a learned behavior, as there appears to be a difference in alternation between bilinguals brought up in "one person/one language" homes and those who have grown up in homes and/or communities where members switch freely between two languages. In many bilingual situations throughout the world, codeswitching has become the norm. In those cases, codeswitching is a general marker of belonging to a mixed group with a multiple identity.

Cross-linguistic influence. Another phenomenon of bilingualism is cross-linguistic influence, or the influence of one language on the other such that features of language A are used while speaking or writing language B. Because not all linguistic features work across languages (i.e., cross-linguistically), this type of influence may result in inaccurate or nonstandard

language. Such influence is also called transfer or interference, although the latter term is considered to be negative and pejorative and reflective of a monolingual perspective of bilingualism. Cross-linguistic application does not indicate confusion on the part of the learner. Rather, it represents worthwhile attempts at communicating and linguistic creativity in the journey of bilingual/second language acquisition and development.

The Advantages of Bilingualism

Current research suggests that there are several overlapping and interacting benefits associated with bilingualism. These benefits encompass communicative, cultural, and cognitive advantages. Sometimes, some social and individual problems are falsely attributed to bilingualism.

Communicative Advantages

Parental Relationships. Bilingual children of multilingual homes are able to communicate in each parent's preferred or dominant language. This may facilitate an optimally close relationship with each parent. In other cases, bilingual children may communicate with their parents in one language and with members of their community or circle of friends in another language. For these families, the native language provides a natural and expressive way to communicate with full intimacy while at the same time providing a link to their past and heritage.

Extended Family Relationships. Bilingualism serves as an intergenerational bridge. Bilingual children who have immigrated to a new country have the advantage of carrying forward the family's heritage language. The monolingual without the heritage family language may feel a sense of alienation from relatives and from the past. When extended family members in another country or region speak only the family's heritage language, bilingualism bridges the generation gap and allows closer relationships, a sense of belonging, and rootedness to develop within the extended family. In this sense, bilingualism contributes to the feeling of family continuity across generations.

Community Relationships. Bilingualism allows for a wider range of communication across people, communities, and beyond. Bilingual children in linguistically diverse communities who study in two-way immersion, or dual-language, programs are able to communicate in the wider community, with school and neighborhood friends, and across continents through Internet communication with pen pals in other regions who speak the target language.

Transnational Communication. Language is one barrier between nations and ethnic groups that sometimes hinders communication and the development of friendly relationships. Bilinguals can act as links, not only within the nuclear and extended family and the community but also across societies. When bilingual people travel within multilingual countries or across borders, their bilingualism allows them to forge relationships with people of different nationalities and ethnic groups that might otherwise have been difficult to establish.

Language Sensitivity. Bilinguals may be more attuned to the communicative needs of their interlocutors, because they must constantly monitor which language to use in different bilingual and monolingual situations. Research suggests that bilinguals may be more empathetic and patient as listeners than are their monolingual counterparts when communicating with others who may not be very proficient in a language.

Cultural Advantages

Bicultural Experience. Because language and culture are inextricably linked, bilingualism provides the opportunity to experience two or more cultures in very authentic ways. The monolingual also can experience multiple cultures, through travel or neighbors or ethnically and linguistically diverse communities, but only as a passive onlooker. To penetrate different cultures and to participate and become involved in the core of a culture require knowledge of the language of that culture. Furthermore, because two languages provide access to a wider cultural experience, there is often a greater tolerance of cultural diversity.

Languages provide unique systems of behavior; folk sayings, stories, histories, and traditions; ways of meeting and greeting; rituals of birth, marriage, and death; ways of interpreting the world; ideas and beliefs; and ways of thinking. Bilingualism provides the potential to actively penetrate the cultures expressed in both languages.

Economic/Employment Opportunities. Bilingualism further provides potential economic advantages. As the global village rises and international trade increases,

bilinguals and multilinguals may be in a relatively strong position in the race for employment.

Cognitive Advantages

Research suggests that bilinguals may have some cognitive advantages, ranging from creative thinking to faster progress in early cognitive development and greater sensitivity in communication. For example, bilinguals may exhibit greater cognitive flexibility in understanding the relationship between objects and their labels. Because bilinguals have two or more words for many objects and ideas, the link between a word and its concept is usually looser for bilinguals than monolinguals. For bilinguals, having two or more words for particular concepts extends the range of meanings, associations, and images. Thus, bilinguals have the possibility of developing more advanced language awareness and more fluency, flexibility, and elaboration in thinking than monolinguals.

"Problems" of Bilingualism

It should also be noted that bilingualism does not come without its share of controversy. Some social and individual problems continue to be falsely attributed to bilingualism. For example, bilingualism is sometimes blamed when young bilinguals exhibit language or personality problems. However, these "problems" tend to be temporary and/or related to broader issues of language attitudes and monolingual perspectives to language development. Another, equally complex issue is the question of identity. Identity is not a problem for many bilingual individuals, who either navigate both languages and cultures effectively or who identify more closely with one language and culture while still valuing and using the other. But some bilinguals (e.g., immigrants) feel the pressure to assimilate and identify with the new country exclusively. This may mean losing the identity of their native country and, possibly, the native language. Again, this "problem" is associated more with monolingual perspectives of bilingualism (where monolingualism represents the natural or normal case of language development, and only one language and culture, usually the majority or mainstream one, must take precedence) than with bilingualism itself. Like bilingualism, identities change and evolve over time, with varying experiences, interactions, and collaborations within and outside a language and cultural group.

The Development of Bilingualism

Given sufficient motivation and opportunity, all normally developing individuals can learn more than one language. The development of bilingualism in children can occur in various ways. Simultaneous or infant bilingualism refers to a child acquiring two languages at the same time from birth, or at least from a very young age. Simultaneous acquisition is often associated with the "one person, one language" principle. Other children hear one language in the home from birth and come into contact with another language later, in the neighborhood, wider community, or school, or as a result of immigration to a new country. This is termed *consecutive* or *sequential bilingualism*. Age 3 is generally regarded as an approximate borderline between simultaneous and sequential bilingualism. There is a large literature on the specifics of childhood bilingual and second-language acquisition, including both natural acquisition conditions and those that occur formally at school. The following section provides an overview of the developmental issues surrounding simultaneous and sequential bilingualism in children.

Simultaneous Bilingualism

Bilingual acquisition is a common and normal childhood experience. It is estimated that two thirds of the world's children grow up in a bilingual environment. Perhaps the most important insight gained from studies on child bilingualism over the past 25 years is that simultaneous acquisition of two or more languages can be characterized as an instance of multiple first-language acquisition. That is, the development of each of the bilingual's languages proceeds in the same way and leads to the same kind of grammatical competence as in monolingual children. Until recently, it was thought that early bilingual development occurred in three stages, moving from mixing the two languages to partial and eventually full language separation. But this framework turned out not to capture the typical developmental pattern of infant bilingualism. Recent research has demonstrated that children as young as 2 separate rather than mix their languages and acquire multiple grammars separately and autonomously.

Children who are learning two or more languages from birth follow the same general pattern as children who are learning one. For bilinguals and monolinguals alike, language acquisition begins at birth or

perhaps even before. During the first year, children pass through several important stages in their language development. Like their monolingual counterparts, bilingually developing babies engage in different types of vocal play, like cooing and babbling, within this first year through which they learn to coordinate their muscles and try out different sounds. At about 1 year, bilingually developing children assemble a vocabulary comprising elements from both languages. Type, quality, and amount of exposure to each language influence which words and concepts are learned in which language(s), as well as the developing grammar of each language. By about 18 months, many children begin to string two words together to make phrases. From then on, their utterances develop and progressively become more complex. As children learn about the boundaries of words and express themselves with the vocabulary available to them during this stage, they might overextend or underextend the meaning of words, as well as use elements from both languages. This elementary form of codeswitching represents children's practical and purposeful attempts to express themselves and is not an indication of language confusion. Codeswitching at this stage may also occur when a phrase or word is simpler or is easier to pronounce in the other language. Another phenomenon of early bilingual development is the creation of language boundaries. The bilingual child learns to associate each of his languages with certain individuals, contexts, and/or situations. In these early years, bilingual children, like their monolingual counterparts, are learning at a tremendous rate, experimenting with language and defining and refining usage. Amazingly, within the same time frame as it takes monolingual children to learn one language, bilingual children learn two languages and become adept at using them in socially diverse and appropriate ways.

Sequential Bilingualism

Childhood Sequential Bilingualism. As with simultaneous bilingualism, there are a variety of routes to childhood consecutive or sequential bilingualism. Some children learn the language of the home first and then become exposed to a second language in the community or when they enter school. Other children learn a second language as a result of immigration. Sequential bilingual development in children is very similar to simultaneous acquisition of two languages

if learning takes place in a natural environment (as opposed to formal or classroom language learning). Single phrases and set phrases are developed initially; children concentrate on simple grammatical constructions, over- and underextend vocabulary and grammar, and mix structures from their two languages as they attempt to make sense of, apply, and navigate their developing bilingual repertoire.

A child's second-language development depends on many factors, including the amount and quality of exposure to each of his or her languages at home, in school, in the community, and so forth. If the child's contact with one of the languages is minimal, it will affect that child's opportunities to use the language, which in turn influences eventual attainment of, and proficiency in, that language. Although a child can learn a second language at any age, it is important that the introduction of a second language not come at the expense of the native language. For example, an immigrant child to the United States who enrolls in an English-only educational program upon arrival will, in effect, be taught in an unintelligible language. In such subtractive situations, the child's native language skills may not continue to develop properly, and the child may struggle to attain the necessary second-language skills to cope in the classroom and compete with native language speakers. Thus, the development of both languages is affected.

Current research suggests that children from minority language backgrounds succeed better when they are taught through their native language while gradually acquiring the majority language. In this way, the native language is built upon and serves as the foundation for future learning. The process of language acquisition is a long and complex one; children who are acquiring a new language while also learning grade-level academic concepts and skills benefit from developmentally, culturally, and linguistically appropriate and supportive instruction, including the continued development of their native language.

Adult Second-Language Acquisition. Adults, like children, develop bilingualism in different ways, contexts, and under different circumstances. In summarizing a framework of second-language acquisition, Rod Ellis suggests that there are five interrelated factors affecting the acquisition of a second language: situational factors, input, learner differences, learner processes, and linguistic output. Situational factors refer to the environment of linguistic interactions (e.g.,

classroom context, formal situation, naturalistic setting), the relationship of the interlocutors, and the topic of the conversation. Factors related to linguistic input concern the type of second-language input received (and its comprehensibility) when listening or reading in a second language. Individual learner differences include aptitude, cognitive style, learning strategies, personality variables, and age. Learner processes refer to those learning strategies through which learners consciously and subconsciously process second-language input, as well as production and communication strategies. Finally, second-language output relates to opportunities to engage in meaningful oral and written exchanges with other language users.

Age is a much debated theme in second-language acquisition and development and thus deserves special attention here. Although important questions still remain unanswered, the evidence gathered so far suggests that the human language capacity is subject to maturational changes. Traditionally, the period between 2 years of age and puberty was regarded as the critical period for the development of first and additional languages, although more recent reviews of the research have largely discredited this notion. The concept of a critical period of language development does not mean that language acquisition is not possible beyond the critical period; rather, it suggests that older language learners resort to other cognitive capacities to develop a knowledge system about a new language. In fact, studies have shown that older children and adults can learn a language more quickly and efficiently than younger children can, because their cognitive skills are more developed and they have a more mature linguistic system to build on. However, it does appear that there are advantageous periods for second-language acquisition. One area that may be affected by age is pronunciation. Some studies have demonstrated that native-like pronunciation might be more difficult to attain for languages acquired after puberty, whereas others suggest that the ability to pronounce a second language in a native-like way is not necessarily lost in adulthood. It may be that pronunciation problems of older second-language learners are attributable to difficulty of access to the target language rather than to an outright loss of the relevant underlying capacity to learn additional languages. Although it continues to be a constant source of controversy, the critical period concept has been a valuable construct in the investigation of second-language acquisition across different age groups.

Conclusion

Bilingualism continues to thrive in a world economy with increased ease of international communications. Although an exact definition of who is or is not bilingual remains elusive and ultimately impossible, the concept of bilingualism has broadened thanks to the growing research base on the topic. Much has been learned about the complex social, psychological, cultural, and linguistic factors and processes associated with becoming and being bilingual. Perhaps most importantly, it has been learned that bilingualism is a common, natural, and dynamic phenomenon that will continue to intrigue researchers and the general public alike for years to come.

Mileidis Gort

See also Bilingual Education; English as a Second Language

Further Readings

Baker, C. (2006). *Foundations of bilingual education and bilingualism* (4th ed.). Clevedon, UK: Multilingual Matters.

Baker, C., & Prys Jones, S. (1998). *Encyclopedia of bilingualism and bilingual education.* Clevedon, UK: Multilingual Matters.

Bhatia, T. K., & Ritchies, W. C. (2006). *The handbook of bilingualism.* Malden, MA: Blackwell.

de Groot, A. M. B., & Kroll, J. F. (Eds.). (1997). *Tutorials in bilingualism: Psycholinguistic perspectives.* Mahwah, NJ: Lawrence Erlbaum.

De Houwer, A. (1990). *The acquisition of two languages from birth: A case study.* Cambridge, UK: Cambridge University Press.

Ellis, R. (1985). *Understanding second language acquisition.* Oxford, UK: Oxford University Press.

Genesee, F. (1989). Early bilingual development: One language or two? *Journal of Child Language, 16,* 161–179.

Grosjean, F. (1985). The bilingual as a competent but specific speaker-hearer. *Journal of Multilingual and Multicultural Development, 6,* 467–477.

Grosjean, F. (1999). Individual bilingualism. In B. Spolsky (Ed.), *Concise encyclopedia of educational linguistics* (pp. 284–290). London: Elsevier.

Hakuta, K. (1986). *Mirror of language: The debate on bilingualism.* New York: Basic Books.

Hamers, J., & Blanc, M. (1989). *Bilinguality & bilingualism.* Cambridge, UK: Cambridge University Press.

Hoffman, C. (1991). *An introduction to bilingualism.* New York: Longman.

Mackey, W. F. (1962). The description of bilingualism. *Canadian Journal of Linguistics, 7,* 51–85.

Romaine, S. (1995). *Bilingualism* (2nd ed.). Oxford, UK: Blackwell.

Wei, L. (Ed.). (2000). *The bilingualism reader.* New York: Routledge.

Bloom's Taxonomy of Educational Objectives

Bloom's Taxonomy is an educational classic. Its formal title is *Taxonomy of Educational Objectives. The Classification of Educational Goals. Handbook I: Cognitive Domain.* A committee of college and university examiners wrote the book, which was edited by Benjamin S. Bloom and published in 1956. Bloom's Taxonomy offered a classification system for educational goals that could be used in the construction of test items and in the formulation of instructional objectives.

The volume presents six categories of educational objectives:

1. Knowledge

2. Comprehension

3. Application

4. Analysis

5. Synthesis

6. Evaluation

Bloom presents sample test items that assess instructional objectives associated with each category of educational objectives along with specific examples of instructional objectives associated with each of the six categories of cognitive objectives.

To Bloom, *knowledge* is defined as "those behaviors and test situations, which emphasize the remembering, either by recognition or recall, of ideas, material, or phenomena" (Bloom, 1956, p. 62). Recognition test items tend to be easier than comparable recall test items, because recognition test items tend to provide incorrect answers along with correct answers (e.g., as with multiple-choice test items). Recall items tend not to provide any answers at all, as the test-taker provides the correct answers to the items.

Bloom clearly acknowledged various forms of knowledge that varied in level of abstractness, from concrete facts to theories and abstract structures. First, there is knowledge of specifics, including terminology and specific facts. Second, there is knowledge of ways and means of dealing with specifics, including conventions, trends and sequences, classifications and categories, criteria, and methodology, Third, there is knowledge of the universals and abstractions in a field, including (a) principles and generalizations and (b) theories and structures. An example of a knowledge objective is "to know the state where New York City is located."

The second category of educational objectives is *comprehension*, which comprises three types of behavior. The first type is translation, which means to transform a communication that is in one form (e.g., a statement in a language, such as English) into a comparable communication that is in another form (e.g., an equivalent statement in a language, such as Spanish). An example of a translation objective is "to graph on graph paper the equation $x^2 + y^2 = 9$." The second type of comprehension behavior is interpretation, which means to formulate a restatement of a communication into another communication. An example of an interpretation objective is "to summarize a short story." The third type of comprehension behavior is extrapolation, which means to make a prediction or to infer an implication or a consequence from a consideration of some statement or communication. An example of an extrapolation objective is "to predict the consequences of implementing a policy of free college tuition for all legal residents."

The third category of educational objectives is application. *Application* refers to using a rule, a principle, or a concept to solve a problem. An example of an application objective is "to solve a distance problem with the use of the Pythagorean theorem."

The fourth category of educational objectives is analysis. To Bloom, *analysis* is "the breakdown of the material into its constituent parts and detection of the relationships of the parts and of the way they are organized" (Bloom, 1956, p. 144). There are three types of analysis behaviors. The first type is the analysis of elements, which refers to breaking down a communication into its constituent parts and identifying the elements of the communication. The second type is the analysis of relationships, which refers to determining the relationships among the elements in a communication. The third type is the analysis of organizational elements, which refers to inferring the organizational principles underlying a communication (e.g., the author's purpose, point of view, or attitude).

An example of an analysis objective is "to infer the view of technology of Mary Shelley in her classic novel *Frankenstein.*"

The fifth category of educational objectives is synthesis. To Bloom, *synthesis* is "putting together of elements and parts so as to form a whole" (Bloom, 1956, p. 162). Synthesis may be viewed as the complement to, and opposite of, analysis.

The products of synthesis behavior fall into three classes: (1) a unique communication, such as a poem or a short story; (2) a plan or proposed set of operations, such as a recipe for how to make use of green pepper slices and pomegranate berries in the same dish; and (3) a set of abstract relations, such as a set of principles for living that interrelate health, a high quality of life, and a limited budget. An example of a synthesis behavior is "to prove the Pythagorean theorem."

The sixth and final category of educational objectives is evaluation. To Bloom, *evaluation* is "defined as the making of judgments about the value, for some purpose, of ideas, works, solutions, methods, material, etc. It involves the use of criteria as well as standards for appraising the extent to which particulars are accurate, effective, economical, or satisfying" (Bloom, 1956, p. 185). The judgments involved in evaluation are based on internal evidence or external criteria. The evaluation of material based on internal evidence addresses features such as "logical accuracy, consistency, and other internal criteria" (Bloom, 1956, p. 188). The evaluation of material based on external criteria makes use of criteria formulated by experts in disciplines or models of excellence in disciplines. An example of an evaluation instructional objective is "to evaluate critically the view of immigration policy suggested in a newspaper editorial."

Knowledge and comprehension are lower-level cognitive objectives, and application, analysis, synthesis, and evaluation are higher-level cognitive objectives. These cognitive objectives form a hierarchy, with knowledge being the objective prerequisite to all other cognitive objectives and evaluation being the objective to which all other cognitive objectives are prerequisite.

Bloom's Taxonomy has informed the writing of achievement test items and the design of instructional programs for decades. It has also inspired subsequent commentaries such as volumes by Lorin W. Anderson and Lauren A. Sosniak, Anderson and David Krathwohl, and Robert Marzano and Thomas Guskey. Bloom's Taxonomy will likely continue to inform educational practice and research in educational psychology.

William M. Bart

See also Assessment; Curriculum Development; Evaluation; Goals; Testing

Further Readings

Anderson, L., & Krathwohl, D. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives.* New York: Longman.

Anderson, L., & Sosniak L. (1994). *Bloom's taxonomy: A forty-year retrospective. Ninety-third yearbook of the National Society for the Study of Education: Part II.* Chicago: National Society for the Study of Education.

Bloom, B. (Ed.). (1956). *Taxonomy of educational objectives, The classification of educational goals, Handbook 1: Cognitive domain.* New York: David McKay.

Marzano, R., & Guskey, T. (2006). *Designing a new taxonomy of educational objectives.* New York: Crown Press.

BRAIN-RELEVANT EDUCATION

Brain-relevant education refers to a relatively new perspective on learning influenced by an abundance of new data emerging from the study of the brain. The term itself—*brain-relevant education*, sometimes referred to as *brain-based learning*—garners criticism by some, who contend that all education relates to, or is based in, the brain. From this perspective, the term may seem pretentious or irrelevant. To others, however, and perhaps the majority, *brain-relevant education* connotes a biological approach to learning. This contrasts with a more traditional position favoring psychological or behavioral emphases.

Other synonymous terms include *education-related brain science*, *educational brain science*, and *educational neuroscience*. Each draws attention to how the brain learns but with particular focus on the biology of learning. This entry uses these terms interchangeably.

Still other similar terms—*neuropsychology* and *biological psychology*—relate neuroscience with psychology more specifically. These subsets of psychology also deal with behavioral oddities and complications resulting from abnormalities in brain structure and function as well as conditioned responses.

The following sections provide a metaview on evolving interest in brain study, relevance of brain science to education, historical background, pertinent philosophical context, themes emerging from brain-relevant education, and pedagogical implications issuing from studies of the brain.

Focus on the Brain

Proclaimed as the decade of the brain by the U.S. government, the 1990s played a significant role in drawing attention to the *science* of cognition. New technology allowed study of the brain in ways not available or practical previously. This technology now makes it possible to see the brain as it functions—an advantage not available in earlier years when study of the human brain was limited to cadaver dissection, on the one hand, and speculation based on interpretation of behavior, on the other.

Methods now commonly used to study live brain anatomy and function include the following techniques and technologies as described by James Kalat.

Imaging of Brain Structure

Computerized Axial Tomography (CT or CAT scan)

In this technique dye injected into the bloodstream creates viewable contrast as x-rays pass through the brain during the scanning process. Resulting photographs can depict structural abnormalities. As this technique utilizes dye injection for contrast, it is classified as invasive.

Magnetic Resonance Imaging (MRI)

This technique applies a strong magnetic field—25,000 times stronger than Earth's magnetic field—in short, repetitive, staccato-like spurts as the subject being imaged passes through a circular tube. Atoms respond to the on–off spurts of magnetism in ways that allow the technology to create images of structures in the brain.

Imaging of Brain Structure and Activity

Electroencephalography (EEG)

By placing electrodes at various locations on the scalp, this technology records electrical activity within the brain. Physicians often provide a sensory stimulus—termed *evoked potentials* or *evoked responses*—to record differences in brain function as a result of the stimulation. This imaging provides approximate location of activity—within about a centimeter of space.

Magnetoencephalography (MEG)

Though similar to EEG, this technique differs in that it measures magnetic fields produced faintly in brain function. This process allows the advantage of observing a given effect in the brain as it travels—somewhat as a wave does on water.

Positron-Emission Tomography (PET)

This technique offers the advantage of high-definition imaging with the added feature of showing hotspots of activity in a live brain. This technique is more invasive that some other brain imaging technologies because it involves injecting chemicals containing radioactive atoms. This fact imposes limitations on the frequency of using this procedure.

Functional Magnetic Resonance Imaging (fMRI)

Similar to MRI, this technique differs in that it detects brain activity based on the presence of oxygen carried by hemoglobin. Both structural and functional imaging result from this procedure.

Fascination with the brain continues well beyond the initial interest stimulated by data revealed with these revolutionary imaging techniques. This is evidenced in how often major news media feature books, articles, and information on the brain. As indicated by attendance at professional conference sessions on the brain and learning and sales of books on this topic, the brain also intrigues educators.

A Question of Brain Relevance in Education

This barrage of brain-related information offers valuable data on the structure and function of the brain. But, is this information relevant to educational practice? Some question the legitimacy of linking brain science to education. John T. Bruer, president of the James S. McDonnell Foundation, is one such individual. The span between brain science and education requires "a

bridge too far," he suggests. Bruer bases this judgment on the risk of inappropriate application of information. A little knowledge about a highly complex organ can easily be oversimplified or misused, especially in the current climate of eagerness and intrigue with the brain.

Others assume a different stance, however. They argue that of all professions, teachers shape brains more significantly than does any other profession. This shaping of the future, they contend, merits legitimacy for teachers to better understand the organ they purposefully influence. Accordingly, many books and much focus on translating neuroscience for educators now exist to guide teachers.

One example is the work of Case Western Reserve University's James Zull. In his book *The Art of Changing the Brain: Enriching the Practice of Teaching by Exploring the Biology of Learning*, Zull explains how neural plasticity—the brain's ability to change and grow with experience—is directly influenced by the quality of what goes on in the classroom. He illustrates how the dendrites of neuron cells in the brain show growth after just 15 minutes of cognitive stimulation. Zull links brain processing to learning theory and further substantiates the need for teachers to gain an understanding of the science of cognition.

A metaview such as Zull's aids in constructing a base of knowledge to promote appropriate applications of science to education. Yet, an inherent question merits consideration: Who determines which authors are reputable for providing guidance and relating this type of information to classroom practice?

The absence of brain science curricula in teacher education creates possibilities for "educational hucksters" to capitalize on this lack of knowledge, according to Robert Sylwester. Early in the emergence of educational brain science focus, Sylwester, one of the first university professors to provide teacher guidance in this area, drew attention to the need for scholarly focus on education-related brain science.

Citing previous vulnerability to "pseudoscientific fads, generalizations, and programs," such as right brain/left brain focus that exaggerated the conclusions of research findings, Sylwester cautioned against potential misuse of scientific data. Without sufficient training of teachers to prudently align brain science with classroom practice, he warned, the general public may become as well or better informed on this topic than are teachers themselves.

In spite of this and similar counsel, teacher education programs have not readily moved to include brain-relevant curricula in teacher training. A major reason for this is the lack of instructors trained to appropriately merge brain science with education. The next section describes progress toward this goal.

Historical Background

University of California, Berkeley

Marian Diamond, from the University of California, Berkeley (UCB), was one of the first pioneers in merging brain science with education. Though a teaching professor at UCB, Diamond is primarily a scientist—a world-renowned neuroanatomist—more than a trained educator. Even so, in the early 1980s at UCB's Lawrence Hall of Science, she and colleagues developed a massive interactive museum display on the brain to educate students and teachers. Busloads of children from California and beyond traveled to this site to learn more about the most complex entity in the universe—the human brain. To further extend this educational support, Diamond and Janet Hopson later wrote a book titled *Magic Trees of the Mind: How to Nurture Your Child's Intelligence, Creativity, and Healthy Emotions From Birth Through Adolescence.*

Association for Supervision and Curriculum Development

Influence of Diamond and others, along with technological advances in neuroscience, continued to make knowledge about the brain more comprehensible to the general public. Subsequently, both high interest in brain science and a perceived need to relate this new information to classroom practice provided impetus for new direction in professional development for education. The Association for Supervision and Curriculum Development (ASCD), a professional organization for teachers and school administrators, turned its attention to the brain and learning.

At ASCD conventions, sessions on education and the brain soared in popularity. In response, ASCD published related books. One of their first publications was Renate Caine and Geoffrey Caine's *Making Connections: Teaching and the Human Brain*, published in 1991. About 6 years later, another of ASCD's published books became a bestseller and remained so for many years—Eric Jensen's *Teaching With the Brain in Mind.* This book maintained a high profile of

popularity with K–12 teachers because it made information about the brain understandable for those with little background in neuroscience. Others, however, responded less enthusiastically to this book and drew attention to its lack of research references supporting claims and inferences. Still, this book played a seminal role in informing teachers about basic concepts of brain science and providing ideas for creating brain-friendly classrooms.

Learning Brain Expo

Keen to the urgency and need for more information for classroom practice, Jensen produced other books, provided intensive teacher training sessions, and organized professional conventions called *Learning Brain Expo*. These 4-day conferences in various locations throughout the United States and beyond continue to attract teachers' attention today. Featuring presentations by scholarly researchers as well as effective classroom practitioners, the conferences focus on classroom efficacy.

Harvard University's Graduate School of Education

About the same time Jensen's efforts gained popularity among educators, Harvard University's Graduate School of Education (HUGSE) began a program called *Mind, Brain and Education* (MBE). To extend the mission of this initiative, HUGSE invited key individuals—from national and international institutes of higher education—to meet at Harvard in groups of 70 for week-long sessions during the summer season. These interactive, collaborative meetings not only provided a certain degree of calibration for educator emphasis on brain science but also spawned a network of professionals to actively promote this focus in teacher training both in North America and abroad. This endeavor heavily emphasized a scholarly, theoretical, and research-based approach to education-related brain science.

Learning and the Brain Conferences

Carrying forward a scholarly emphasis, *Learning and the Brain* conferences developed as a joint venture between Public Information Resources, Inc., the Kosik Neurobiology Lab at Harvard Medical School (now at the University of California, Santa Barbara),

and the MBE program at HUGSE. These conferences feature a distinguished faculty of professors, neuroscientists, and researchers who offer dynamic perspectives on the brain and learning. This venture also provides member benefits in an organization called the *Learning and the Brain Society.*

Whereas the Jensen *Learning Brain Expos* tended to be located on the West Coast of the United States, the Harvard collaboration focused on East Coast offerings—in the Boston/Cambridge, Massachusetts, area or at the National Institutes of Health in the Washington, D.C., area. Currently, both of these conferences convene on both coasts of the United States and beyond.

International Mind, Brain and Education Society

Another professional society also spawned from this Harvard collaboration. Kurt Fischer and Marc Schwartz, along with others affiliated with the HUGSE's MBE program, formalized their affiliations by establishing the International Mind, Brain and Education Society (IMBES). Aimed at promoting research and scholarly dialogue, this organization offers a refereed journal and newsletter, book reviews, conferences, and other member benefits.

A metaview on brain-relevant education suggests that a tension exists between those whose main focus is research/psychology/learning theory and those who emphasize practical aspects of educational brain science. On the premise that disequilibration is vital in warding off rigidity and stasis—not just in biology, but in education as well—this tension between theory and practice provides benefit in promoting active dialogue and forward progress. The next section offers additional rationale for this perspective.

Philosophical Metaperspective

Many voices speak about the brain and learning—from macro- and micro-levels of interpretation and from holistic perspective. Philosophical tensions have much to do with what these voices say.

Macro-intersystemic voices tend to describe brain activity as a phenomenon of its environment. They speak of inferred brain activity based on observable evidence of brain activity in behavior and brain malfunction, movement, and language.

Metacognition and abstract analysis typify this perspective on the brain. Proponents of this perspective often describe the brain in cause-and-effect terms and metaphoric analogies such as computer hardware and software. Voices in this category tend to be those of cognitive psychologists, linguists, physical anthropologists, philosophers, artificial intelligence experts, and neuropsychologists.

On the other hand, *micro-intrasystemic voices* reflect on the brain by studying brain activity from within the human unit. They describe cells, molecules, amino acids, neurochemicals, and smaller cellular systems within more complex systems. This perspective focuses heavily on quantification or measurement of brain activity. Voices in this category tend to be those of neuroscientists, neuroanatomists, neurophysiologists, and neurochemists.

The graphic in Figure 1 depicts the educator's relationship to these two categories of specialists.

Voices on the Brain

Cognitive Psychologists
Linguists
Physical Anthropologists
Philosophers
Artificial Intelligence Experts

Educators → ← Educators

Neuroscientists
Neuroanatomists
Neurophysiologists
Neurochemists
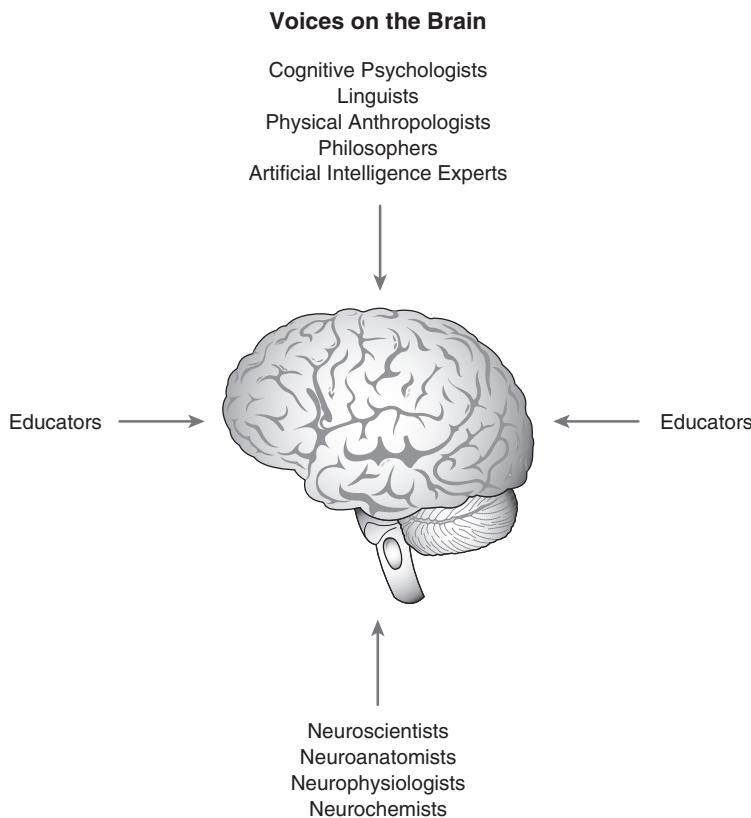
Figure 1   Voices on the Brain

Source: Copyright 2001 Linda Bryant Caviness. No duplication without permission.

Note: This graph depicts varying perspectives on the brain.

Typically, educators are not trained as specialists in behavioral or biological sciences. Yet, societal pressures call for more focus on the brain in education—or the science of education. To date, however, state and federal departments of education have not taken action to infuse teacher training with this burgeoning body of new information.

Reductionist Versus Constructivist

Added to this lack of professional training is another layer of involvement—one implied previously in regard to the tension between two foci: theory/ research and classroom practice. Much of the tension here results from two varying perspectives—one favoring a reduced, fragmented, controlled, linear, measurable emphasis and the other honoring a holistic, qualitative, naturally represented reality. Though seldom articulated, this tension is strong in the realm of education-related brain science.

A Newtonian–Cartesian mind-set—the philosophical perspective that tends to admit to reality only measurable evidence that has undergone the process of reductionistic examination—continues to be the heavyweight champion today in the sciences and in educational psychology and philosophy. The No Child Left Behind (NCLB) initiative demonstrates the strength of this influence. Before NCLB, constructivist processing and learning commanded a respectable position in bureaucratic function. However, NCLB quickly returned the pendulum's swing to a more controlled and quantitative position on assessment and instruction.

This position promotes the value of direct instruction and controlled curricula, as opposed to nurturing creativity and allowing the student to share the role of the teacher and to be in charge of his or her own learning at least some of the time. According to Zull, movement is cognition expressed, and that movement is most meaningful cognitively when it stems from intrinsic drive—a personal desire from within to learn.

Proponents of brain-relevant education recognize the value of both quantitative and qualitative perspectives in research to

better understand the brain and learning. They value behaviorist and constructivist contexts. Too much weight on either of these dualistic positions diminishes valuable tension that prevents debilitating stasis. This same need for equalized tension represents itself in the following relationships with regard to the brain and learning:

- Nature Versus Nurture
- Passive Reflection Versus Active Processing
- Extrinsic Drive Versus Intrinsic Motivation
- Concrete Thought Versus Abstract Thought
- Skills Versus Creativity
- Objectivity Versus Subjectivity
- Control Versus Freedom
- Specialization Versus Holism

Education-Related Brain Science Themes

A conclusive finding in the study of the brain reveals that the brain functions symbiotically. Linear processing certainly occurs throughout the brain, but many of these linear processes happen simultaneously. Interconnectivity defines brain function.

This principle reveals itself in themes surfacing in brain-relevant education. These emerging themes suggest that learning is best facilitated with holistic nurture that equally honors mental, physical, and emotional/social/spiritual stimulation. These themes include, but are not limited to, the following:

Body and Mind. Body and mind are integrally related in all life processes, including the ability to learn, and the heart is surfacing as a significant player in this dynamic relationship.

Senses. Learning is a multisensorial, integrated process, which involves the body and brain, not just the brain alone.

Exercise and Movement. Physical exercise and movement empower learning by promoting oxygen exchange and vitalization of various physiological processes.

Health Habits. Health habits (nutrition, fresh air, sunshine, exercise, water intake, rest, low stress levels, moderation, and regularity) profoundly influence the brain, its development, and its ability to learn.

Emotions and Neurochemistry. Emotions and neurochemistry facilitate and transact mental processes and bodily functions as they aid information transfer and create mental states through repetitious disequilibrations—moving from balance to imbalance repeatedly by means of electrochemical exchange.

Music and the Arts. Music and the arts facilitate brain development and maintenance through highly integrative stimulation.

Attention. The brain's attention—how it is gained and maintained—is driven by sensorial intake, physiological factors, levels of motivation, memories, and degree of novelty perceived in relation to prior knowledge.

Social Influences. Sensory input from community and environment promotes brain development and stimulates the expression of genetic proclivity.

Plasticity and Enrichment. Areas of the brain are plastic and capable of changing or growing with varying levels of nurture and environmental influence.

Stages of Development. Brain development stages create optimal times—often referred to as critical periods—for particular learning opportunities.

Language. Language shapes the brain, thinking, and culture.

Making Meaning. Learning involves making connections within complex neural and chemical networks, and relating new knowledge to prior knowledge is how the brain constructs meaning.

Individualism. The learner is a complex, unique individual, and each brain is uniquely different and functional.

Motivation. When intrinsic motivation drives the learner, the brain can and does create its own rewards; however, though extrinsic motivation tends to promote skills development, it can lessen creativity if the brain is not allowed also to be in charge of its own learning.

Memory. Rather than being a discrete, singular skill in a particular area of the brain, memory is a process subject to constant neurochemical change as a result of changing stimuli and emotional states; furthermore, memory can develop in various locations and organs within the brain and body.

Theme Connectivity

Figure 2 illustrates further the symbiotic nature of brain function as mentioned previously. Placing each of the themes (listed previously) around the exterior of a circle and drawing lines from each theme to other connected themes demonstrates
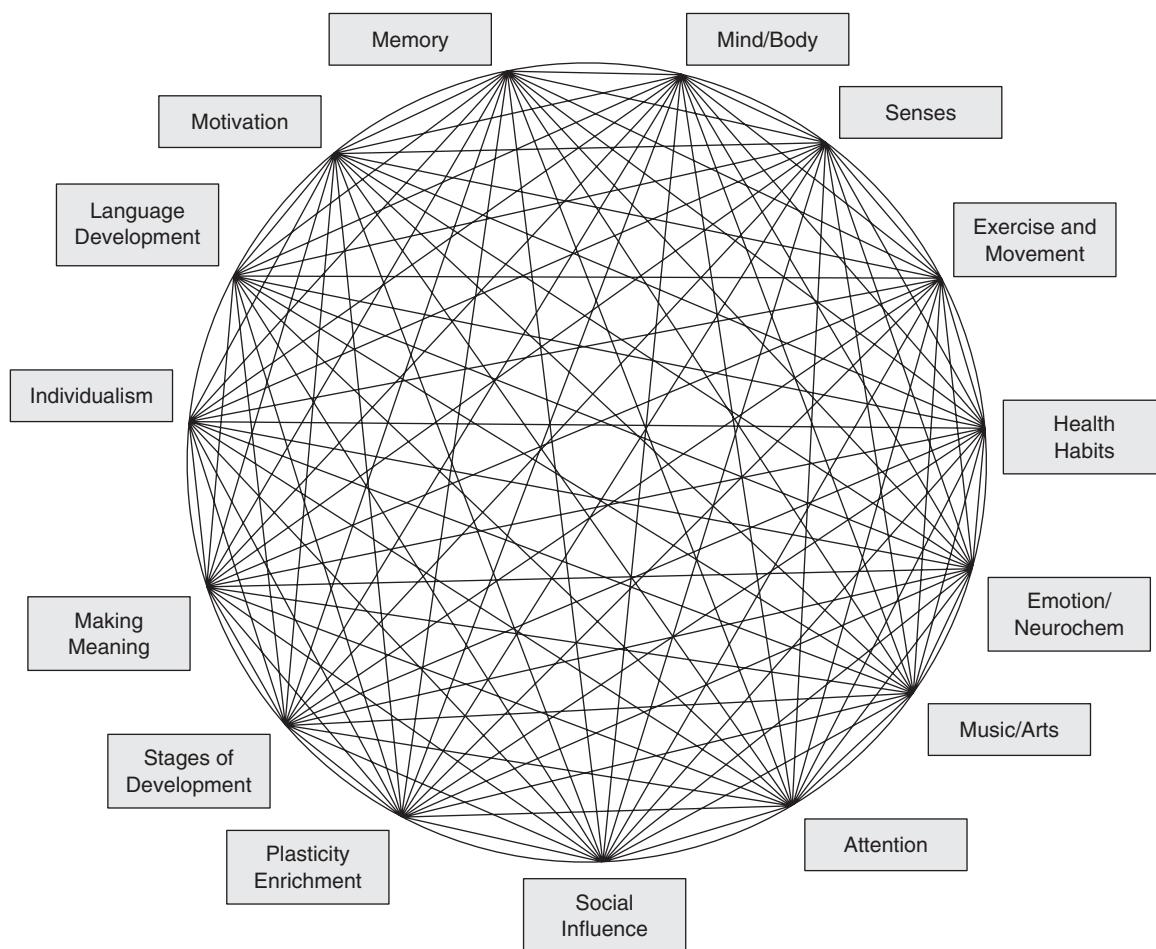
Figure 2   Theme Connectivity

integral connectivity—much like the dynamic connectivity now known to exist among cells, systems, and subsystems of the brain itself. This interrelatedness may provide insight for better understanding of the importance of creating multidisciplinary and integrated learning environments.

Holistic Brain Nurture

Throughout much of the professional literature on brain-relevant education, writers reiterate a major concept: The brain thrives with nurture that honors balance among mental, physical, and social/emotional/spiritual stimulation. Linda Bryant Caviness postulates that this triad emphasis of mental, physical, and spiritual function actually defines the structure and physiological

operation of the brain and body and its constituent parts in fractal-like fashion. Mental, physical, and spiritual functions together constitute a whole. To separate out one aspect of the human wholeness as more important than others disregards the integrity that defines the human brain and distinguishes it above other animal forms.

Traditionally, however, education has favored emphasis on mental nurture. Schools and school systems often cut physical education and the arts first when budgets undergo cost reduction. Though education often claims to honor holistic mental, physical, and spiritual nurturing, in actuality a philosophical perspective that favors reduced, fragmented, linear conceptualization drives the decisions that shape pedagogical practice. Brain-relevant education exposes this dualistic reality.

The science of heart–brain connections now contributes to brain-relevant education. This research reveals an intimate relationship between heart and brain. Now described as a sensory organ with a little brain of its own, the heart significantly influences the quality of cognition. Recognizing the need to better understand this dynamic relationship, Ohio's Cleveland Clinic recently received $17.3 million to establish the Earl and Doris Bakken Heart-Brain Research Institute. These new data further contribute to the importance of education that is holistic.

Brain-Friendly Teaching and Learning

Howard Gardner's multiple intelligences, Daniel Goleman's emotional intelligence, Mihalyi Csikszentmihalyi's flow concept, and others all contribute to current knowledge about teaching and learning. Educators in the classroom, however, seek a particular model or schema to guide them while under the pressure of classroom practice. They desire a plan that combines the advantages of multiple learning theories as well as the knowledge gained from study of the brain.

One model that closely matches these expectations is Bernice McCarthy's 4MAT model for teaching and learning. This easy-to-remember plan is designed to honor three major tensions: active and passive processing, concrete and abstract conceptualization, and extrinsic and intrinsic motivation. Four major steps help teachers hold themselves accountable in respecting brain-friendly learning.

Step 1 helps students connect with prior knowledge and to examine what they already know or don't know. Here, the teacher engages student interest by fostering curiosity.

Step 2 provides opportunity for teachers to model and define the new concept didactically. Visual aids, demonstrations, and textual exposure help to introduce this content.

Step 3 allows students to physically experiment with the new concept under teacher guidance. This prevents internalization of erroneous information. This step also provides time for students to begin to extend the new concept to their own context.

Step 4 encourages students to refine their extended application through peer feedback before they teach their creative extension to others. As teachers swap roles with the students, they can easily determine how well students internalized the new concept.

McCarthy's model aligns well with Zull's explanation of how information is processed throughout the brain. Sensory data travel from brain stem to limbic area and on to the cortex. In the cortex of the brain, where conscious thought or higher order thinking takes place, information moves from sensory integration to conceptual integration and finally to motor integration. Similarly, McCarthy's model connects with the student's context socially/emotionally/spiritually, then promotes new conceptualization, and finally involves the student physically and creatively in real-life applications.

Knowledge about teaching and learning continues to develop. Brain-relevant education has only begun to add to the base of knowledge accumulated over many years of focus on behavior. In good faith it will continue to expand our understanding of human potential and the challenge of its nurture.

Linda Bryant Caviness

See also Malnutrition and Development; Maturation; Neuroscience

Further Readings

Bohning, D. E., Lorberbaum, J. P., Shastri, A., Nahas, Z., & George, M. S. (n.d.). *Structural brain imaging (CT and MRI) in primary psychiatry.* Retrieved February 15, 2007, from http://www.musc.edu/fnrd/primer_struct.htm

Bruer, J. T. (1997). Education and the brain: A bridge too far. *Educational Researcher*, 26(8), 4–16.

Caine, R. N., & Caine, G. (1991). *Making connections: Teaching and the human brain.* Alexandria, VA: Association for Supervision and Curriculum Development.

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience.* New York: Harper & Row.

Diamond, M. C., & Hopson, J. (1998). *Magic trees of the mind: How to nurture your child's intelligence, creativity, and healthy emotions from birth through adolescence.* New York: Penguin Putnam.

Gardner, H. (2006). *Multiple intelligences: New horizons.* New York: Basic Books.

Goleman, D. (2006). *Emotional intelligence: Why it can matter more than I.Q.* (10th anniversary ed.). New York: Bantam Books.

Jensen, E. (1998). *Teaching with the brain in mind.* Alexandria, VA: Association for Supervision and Curriculum Development.

Kalat, J. W. (2007). *Biological psychology.* Belmont, CA: Thomson Wadsworth.

McCarthy, B. (2000). *About learning* (2nd ed.). Wauconda, IL: About Learning.

McCarthy, B. (2006). *About teaching.* Thousand Oaks, CA: Corwin Press.

Sylwester, R. (2003). *A biological brain in a cultural classroom: Enhancing cognitive and social development through collaborative classroom management.* Thousand Oaks, CA: Corwin Press.

Sylwester, R. (1995). *A celebration of neurons.* Alexandria, VA: Association for Supervision and Curriculum Development.

Zull, J. E. (2002). *The art of changing the brain: Enriching the practice of teaching by exploring the biology of learning.* Sterling, VA: Stylus.

BULLYING

Bullying can be defined as consistent, purposeful negative behavior that is directed toward another individual or the persistent abuse of power, which is perpetrated by children and adolescents against their more vulnerable peers. It can take many forms including physical, verbal, and social aggression and is a prevalent part of school life that can have serious, detrimental effects on those who are involved. It can also occur among adults in workplace settings. Although girls and boys are equally likely to be involved in bullying, girls are more likely than boys to use social aggression. Bullies themselves are often quick tempered and may target their victims simply for fun or because they often have many witnesses, including teachers, and can gain a great deal of status from engaging in this type of behavior. This implicit sanctioning of bullying helps to create an unsafe environment. Intervention and prevention programs that are targeted at the whole school are considered to be the most effective in reducing levels of bullying.

Bullying includes physical aggression (e.g., hitting, kicking, and aggressive gestures) and verbal aggression (e.g., threats, insults, and mocking). Other forms of social manipulation that are directed at undermining the individual's social standing (e.g., spreading distasteful rumors) can also constitute a form of bullying.

Bullying is distinguished from more general aggression by the nature of the relationship between the bully (also referred to as the perpetrator) and the target (also referred to as the victim). In this relationship, the bully is usually seen as a more powerful individual or as part of a more powerful group. The victim, in contrast, is usually seen as weaker or different and hence vulnerable. Power and vulnerability in this context does not necessarily refer to physical strength but can refer to a number of physical and personal attributes.

Bullying has most often been studied and considered within the school context, although recently there has been interest in workplace bullying. Thus, the issue of bullying is relevant to educational psychologists working in school and vocational settings. In the school context, the educational psychologist has the responsibility of ensuring that children are safe at school and that bullying behavior is detected, reduced, and prevented. Ensuring the safety of children and young adults is important not only from a legal standpoint but also from a humanist standpoint. According to Abraham Maslow, two of our basic human needs are safety and a sense of belonging or connection. If an individual is threatened by a bully, then his or her sense of safety can be undermined. Furthermore, victims of bullying are unlikely to feel connected to their school, classroom, or peers. Without this sense of belonging or connection, victims of bullying will have difficulty focusing their attention on learning.

The educational psychologist can work with colleagues to ensure all personnel and students are aware that bullying is unacceptable and are able to identify and report instances of bullying. Part of the role of the educational psychologist is to ensure a school climate in which bullying is not acceptable. When bullying has been detected or reported, the educational psychologist can design interventions to stop bullying by intervening directly on the behavior of both the bully and victim. Strategies to prevent bullying can also be designed and implemented at the whole school level.

A similar array of bullying behaviors that are seen in schools and vocational settings are also found in the workplace. However, workplace bullying often is more subtle than the overt acts of aggression observed in schools. For educational psychologists who are working with organizations, a similar set of principles should underpin their ability to intervene and prevent this sort of behavior; a detailed account of workplace bullying is beyond the scope of this entry, however.

It is important to distinguish bullying from other types of negative or aggressive behavior that is directed toward others. In addition to the power imbalance, bullying, by definition, refers to behaviors that are intentional, are persistent, and may involve an accumulation of incidents. Although violent or hostile behavior is similar to bullying, it can be a one-time event, and the

victims may be randomly chosen. Bullying, in contrast, is characterized by a persistent pattern of ongoing negative behavior directed at one or more specific targets or victims.

Harassment and *teasing* are other terms that are related to bullying. The former is often used with reference to sexual harassment; similar to other bullying behavior, harassment frequently is an ongoing behavior acted out against a particular target. When referring to the school context, sexual harassment can be subsumed under the heading of gender-based bullying. Teasing refers to a broad range of behaviors. At one end of the spectrum, teasing can be light-hearted banter between friends; however, the nature of the relationship between the individuals usually distinguishes teasing from bullying. If there is a clear power differential, the line between teasing and bullying can become unclear, especially when the teasing is persistent, one-way, and akin to mockery.

Forms of Bullying

The most obvious forms of bullying behavior are those that involve physical aggression, which may be minor or may lead to severe physical injuries or disfigurement. As with any sustained assault, bullying may result in psychological trauma for the victim. Other forms of bullying behavior may be verbal in nature and can include name-calling, threats of physical aggression, and hostile teasing. Similar to physical bullying, verbal bullying also has the potential to have a negative psychological impact on the victim. The final category of bullying behavior includes more indirect forms that are designed to socially exclude the victim or damage the individual's social standing by changing the peer group's perception of the victim. These forms of behavior may be overt and include negative body language such as turning away, staring, and eye rolling. They may also include verbal or written comments that ostracize the individual from his or her social world. This destructive social behavior can include the manipulating friendship groups, publicly sharing confidential information, or spreading unpleasant rumors so that the individual is socially excluded from his or her peers. Some aspects of this later form of social aggression have also been referred to as *relational aggression.*

It is important to note that often an individual may be subjected to both overt and covert forms of bullying, and the lines between the two types can become blurred, particularly when the behavior is not physical. Furthermore, if the bullying is covert, the perpetrator may remain unknown to the victim; this situation can result in further psychological trauma for the victim.

Bullying in Context

Physical and verbal bullying behavior can occur on the playground, lunchroom, before and after school (e.g., on the bus or near the school entrance), at school camps, or during school excursions. For high school students, the bullying may occur in hallways during the transition between classes. Although bullying usually occurs out of the teacher's sight, more subtle but direct behaviors may occur within classrooms in front of teachers. These behaviors may include social exclusion and the use of negative body language or other comments and actions. In the presence of the teacher some of these may be disguised as "teasing" or as "an accident" (e.g., tripping an individual, leaving something unpleasant on a child's seat).

Over the past decade, the nature of bullying has begun to change with the increased use of technology. For example, computers and the use of the Internet are a common part of school life in most developed countries. The use of Internet connections at home and the increased use of mobile phones by children and adolescents (including text-messaging) have resulted in targets being more accessible to bullies virtually 24 hours a day, 7 days a week. Not only are victims more accessible now, but the perpetrators are better able to remain anonymous. Whereas traditionally an anonymous note or phone call would have been the only option, perpetrators are now able to use text messages and unidentifiable e-mail addresses to ensure their anonymity.

Children Who Are Bullied

Victims of bullying are often cautious children who appear vulnerable, anxious, and lacking in confidence. Children also may be targeted simply because they are new to a school or neighborhood, have few friends, and are not obviously a part of any social group. In addition, children may be targeted for any number of seemingly innocuous reasons. For example, they may be wearing the wrong clothes, be overweight, seem weak, or be too tall or too short. Bullies may also focus on an individual's religious affiliation, ethnicity, race, or socioeconomic status. Individuals may also be

bullied because of their sexual orientation. However, there is some evidence to suggest that the cultural and social context will have a significant impact on the relative prevalence of bullying based on these factors. Whether the abuse is directed at a particular individual or not, oftentimes the climate or atmosphere of hostility toward individuals who are GLBT (i.e., gay, lesbian, bisexual, or transgendered) is likely to be apparent within a class or school context because derogatory terms can permeate the classroom language.

Gender-based bullying refers to bullying behavior that is targeted toward an individual because of his or her gender. Although this type of bullying has been common in coeducational schools for some time, it is only recently been considered as a form of bullying. It can be directed at either girls or boys and may be related to any aspect of the individual's appearance, interests, or behaviors that are counter to sex-role expectations. For example, if a girl has a very short haircut or if a boy wears a pastel shirt, each may become a target of gender-based bullying. Also, if a boy shows an interest in a nontraditional occupation, such as dance, or a girl wants to participate in football, each may be bullied about his or her choice. When the bullying behavior has distinctive sexual overtones, it is likely to be classified as a form of sexual harassment. Although this type of bullying can occur to both boys and girls at any stage of their development, it is more likely to occur in adolescence and to girls who have reached puberty earlier than their peers.

Another group of individuals who may be targeted for bullying behavior are those who have some form of disability, whether it is a physical disability, learning difficulty, intellectual disability, or social or emotional disorder. Although the mainstreaming of all individuals with disabilities has had many positive outcomes and is considered, in most circumstances, to be best practice, there have also been some costs. In particular, individuals with disabilities may become targets for bullying, particularly if the school or classroom teacher has not engaged in any diversity training or instruction. For some individuals, particularly those who have intellectual disabilities, not only is their ability to defend themselves limited but they are also less likely to have any allies within the school.

According to a report commissioned by the American Medical Association, some of the strategies victims employ to cope with the bullying include enduring the bullying, reporting the incident to school personnel or a family member, seeking revenge, retaliating, finding allies, or ignoring the bullying. Often, however, the bullying goes unreported; less than a quarter of bullying incidents are reported to an adult, according to the American Medical Association.

Identifying Bullies

Although not all children engage in bullying behavior, the typical bully is not easily distinguishable from peers. That is, boys and girls from all socioeconomic backgrounds, religious affiliations, cultures, sexual orientations, and abilities have the potential to engage in bullying behavior at some point in their lives. However, according to a report written by Dan Olweus and colleagues, it appears that there are a number of individual, peer, family, school, and community factors that contribute to the likelihood that a person will engage in bullying behaviors. In particular, children who are easily angered, are impulsive, view violence in a positive way, and have low levels of empathy toward others are at risk for bullying. Other risk factors include low levels of parental warmth, permissive parenting, and harsh discipline. Within the school environment, peers can play a significant role if they have similar views to the child with regard to the acceptance of violence. Furthermore, a school culture that accepts bullying and fails to closely monitor children's behavior can also be a contributing factor. It is the interaction of these factors that puts a child at risk for engaging in bullying behavior.

There is also evidence to suggest that some individuals may engage in the behavior because they lack social skills. Their social inadequacy results in an inability to form appropriate relationships with the peer group. This deficit in effective social communication can have numerous causes, including modeling from the home environment, where they may witness bullying behavior between siblings or parental figures. However, the social skills deficit model has been questioned because bullies who engage in social aggression appear to have fairly sophisticated understandings of social behavior. Thus, it appears that a lack of social skills may be associated with some but not all bullies.

Contrary to popular belief, bullies are not necessarily social outcasts; in fact, most have friends at school and are very much a part of their peer group. They often have a healthy self-esteem and can at times be part of the "popular" crowd. Although this seems somewhat counterintuitive, given that popular children have traditionally been considered friendly,

cooperative, and well liked, recent research in the field of peer relations has highlighted that there is a subset of popular individuals who are "popular but not well liked." That is, other members of the peer group may hold these popular individuals in high regard, but they also may recognize that these individuals may be very socially aggressive so they do not necessarily make desirable friends. For a small subset of bullies, the cycle of victimization is continuous because they are also victims of bullying. These bully-victims are often quick tempered and try to retaliate in a violent way and may then, in turn, bully younger or weaker children. Similar to victims, bully-victims are also at a high risk for internalizing disorders such as anxiety and depression.

Reasons for Bullying

According to Ken Rigby there have been numerous reasons put forth as to why some children and adolescents engage in bullying behavior, including for fun, for revenge, or because of peer group pressure. The notion that bullying behavior is enjoyable has some support in operant theory, which suggests that bullies engage in this behavior because it is rewarding. The behavior might also be shaped and maintained because of the consequences of bullying. That is, bullies are positively reinforced for their behavior because they find it enjoyable or they are able to gain tangible objects (e.g., money, food, or other desirable items the target may be forced to relinquish). Furthermore, bullies may gain positive reinforcement in the form of attention from their peers, with the possible increase in status within the peer group. It is also possible that bullies may receive some personal satisfaction in intimidating another individual, as it makes them feel more powerful, particularly if they feel powerless or inadequate in other areas of their school or home life. Thus, bullies may engage in the behavior to alleviate boredom, to entertain themselves, or to distract themselves from the more serious and difficult task of schoolwork.

The notion that some children bully because of peer pressure or to entertain their peers highlights that bystanders or witnesses can play a significant part in bullying. Witnesses fall into a number of categories of involvement, including those who (a) do not instigate, but still take part in the bullying, (b) encourage the behavior, (c) passively watch, or (d) overtly discourage it or report the incident to adults. Whatever their involvement, passive witnesses play a vital role,

as they make up the majority of the school environment. By failing to stand up for or defend a victim, peer witnesses are implicitly condoning the behavior and reinforcing the normality of the situation as well as implying that it is somehow the victim's fault. The same applies for adult witnesses who may unwittingly overlook bullying behavior or consider it to be a natural part of growing up and thereby provide tacit support to the bullies.

Consequences of Bullying

Although there is some evidence to suggest that not all children who are victimized suffer long-term effects of their experiences, those who are affected, either as bullies or as victims of bullying, can suffer significant repercussions with regard to their learning outcomes, health, and socioemotional well-being. For example, victims may suffer from low self-esteem, loneliness, anxiety, and/or depression, all of which may influence their immunity to illnesses. Victims are more likely to suffer from stomachaches, headaches, and suicide ideation; in extreme cases they may engage in self-harm or suicide.

When victims' mental health is compromised it can have a significant influence on other aspects of schooling, including their relations with peers, academic performance, and their willingness to engage in extracurricular activities. Victims are unlikely to feel safe, and because they need a strong sense of security to prosper at school, they may disengage from school. In addition, victims may not have the coping skills or social networks to defend themselves, which increases their vulnerability. Victims of bullying may also engage in externalizing behaviors, such as overt aggression, or other risk-taking behaviors, such as increased drug and alcohol consumption. There are also consequences of bullying for the bullies themselves. They are at increased risk for dropping out of school and are more likely to engage in criminal activities as young adults.

Prevalence of Bullying

To determine the extent to which bullying occurs in schools, a range of different measurement procedures have been developed, including teacher reports, peer nominations, phone-ins, interviews, and direct observations. The most commonly used measure of bullying prevalence is through self-reports, which is the method used in the pioneering work of Dan Olweus

and his colleagues in Scandinavia in the 1970s. Olweus's large-scale studies of 150,000 schoolchildren in Grades 1 through 9 found that 15% of children reported being victims of bullying several times during the previous 3- to 5-month period.

It appears that little has changed in the intervening years with similar levels of victimization found in countries such as the United States where a number of surveys have found bullying rates hovering around 20%. However, there appears to be a developmental progression both in the types of bullying used by children and adolescents and in the prevalence of bullying. Overall, the most frequently used form of bullying for all children is verbal bullying. Children become less physical as they get older and begin to engage in more verbal and social forms of bullying. In addition, there is a fairly consistent decrease in the amount of bullying children experience (or at least report) through the elementary grades. According to a report by Tonja Nansel and her colleagues, bullying rates appear to drop from 25% to 10% between the sixth and tenth grades. However, there is evidence to suggest that a resurgence of reported bullying incidents occurs around times of transition, such as when children are moving from elementary to secondary school, where they are once again the youngest and smallest members of the peer group.

Gender and Bullying

Although boys and girls are just as likely to engage in bullying, some differences have been identified in terms of the types and frequency of bullying. In particular, boys tend to bully more frequently than girls, and boys are more likely to engage in physical and verbal bullying behavior than are girls. Girls, in contrast, are more likely than boys to engage in social bullying. In terms of victimization there are conflicting findings, with some studies showing that girls and boys are victimized at similar rates and others suggesting boys are victimized more often than girls. With regard to the targets of bullying, boys tend to bully both girls and boys, whereas boys are usually victimized only by other boys.

Intervention and Prevention

For many years bullying was considered by some adults to be just a natural part of growing up and something that children and adolescents had to

endure. However, over recent decades a sufficient number of international reports highlighting the potential effects of bullying on individuals' mental health—including the risk of suicide—have altered this prevailing attitude. Today there is more recognition of the need to prevent and reduce bullying. To intervene and ideally prevent bullying from occurring, adults in schools need to be acutely aware of the myriad forms of bullying. The accurate identification of bullying has therefore been an important issue for educators and school psychologists. It appears that adults in schools tend to overestimate their ability to intervene and address bullying. Studies have shown marked differences between students' and teachers' perceptions of the effectiveness of adult interventions.

Traditionally, the intervention strategy of choice was to target the bully for punishment or rehabilitation; however, this has been found to be an ineffective method that can backfire on the victim because the perpetrator may retaliate against the victim for reporting the incident. One attempt at rehabilitation is to bring the perpetrators together to talk about their behavior and to teach them alternatives to aggressive behavior. However, the net effect is bullies may gain some sense of power from this acknowledgment and may also learn new bullying techniques from their peers. Along similar lines, the victim has been targeted for remedial work in the area of social skills and coping strategies. Again, this approach alone is usually insufficient to alleviate the problem because it inadvertently places the blame on the victim.

When school systems, policies, and procedures are not cognizant of the detrimental effects of bullying, it is unlikely that the staff and students will take prevention and intervention seriously. Therefore, the most successful methods of prevention and intervention are those based on a comprehensive whole-school approach first proposed by Dan Olweus and colleagues. The Bullying Prevention Program has been used across hundreds of schools worldwide and has proven to be effective in reducing bullying. The basic premise of the program is that bullying behavior has its roots in the culture of the school, and if educators and parents believe that children have a basic human right to feel safe at school, then the school has a responsibility to create that safe environment. This whole-school approach requires a systematic commitment to changing the prevailing culture; to do so, it is necessary to intervene at all levels in school by including the

individual children, classmates, teachers, administration, and parents in all aspects of the program.

Vanessa A. Green

See also Aggression; Conflict; Emotional Intelligence; Psychosocial Development; Risk Factors and Development

Further Readings

American Medical Association Educational Forum on Adolescent Health. *Youth bullying.* Retrieved July 20, 2006, from http://www.ama-assn.org/ama1/pub/upload/mm/39/youthbullying.pdf

Espalage, D. L., & Swearer, S. M. (2004). *Bullying in American schools: A social-ecological perspective on prevention and intervention.* Mahwah, NJ: Lawrence Erlbaum.

Olweus, D. (1993). *Bullying at school: What we know and what we can do.* New York: Blackwell.

Olweus, D., Limber, S., & Mihalic, S. (1999). *The Bullying Prevention Program: Blueprints for violence prevention.* Boulder, CO: Center for the Study and Prevention of Violence.

Rigby, K., & Bagshaw, D. (2003). Prospects of adolescent students collaborating with teachers in addressing issues of bullying and conflict in schools. *Educational Psychology*, 23(5), 535–546.

Smith, P. K., Peplar, D., & Rigby, K. (2004). *Bullying in schools: How successful can interventions be?* New York: Cambridge University Press.

Nansel, T. R., Overpeck, M., Pilla, R. S., Ruan, W. J., Simons-Morton, B., & Scheidt, P. (2001). Bullying behaviors among U.S. youth: Prevalence and association with psychosocial adjustment. *Journal of the American Medical Association*, 285(16), 2094–2100.

C

Creativity is a type of learning process where the teacher and pupil are located in the same individual.

—Arthur Koestler

Calculator Use

Computers, four-function calculators, and graphing calculators have assisted in the development of significant advances in the field of mathematics. For example, the famous 4-color theorem went unsolved for more than 100 years until 1976, when the calculation and visualization powers of computers were utilized to derive a mathematical proof. Advances in production speed, manipulation, and representation of data through the use of calculators and computers have also benefited the field of statistics, including the specific area of educational statistics. Educational psychology has taken advantage of these advancements in product technologies to support research and theory in the field, particularly in regard to these technologies' abilities to support statistical computation and manipulation.

Although the first desktop calculators were available in the 1960s, their commercial availability did not emerge until the early 1970s. These early "four-function" calculators were limited in scope in that they were able to perform only the basic mathematical operations of addition, subtraction, multiplication, and division. In the late 1970s, companies such as Casio, Hewlett-Packard, Sharp, and Texas Instruments began to produce programmable calculators. These models had small memory chips capable of storing a limited number of preprogrammed algorithms, including square root, exponents, and basic statistics. Some of these models also allowed users to input basic computer code and program their own algorithms.

In the early 1990s, graphing calculators greatly expanded the user capabilities of the handheld devices. The ability to compute fixed algorithms, represent functions and data in multiple ways, and support simple computer codes, made the graphing calculator a prime example of a product technology that allows for an enormous amount of mathematical power, speed, and visual representation. Graphing calculators readily support the production and use of educational statistics through the calculation of statistical tools such as mean, standard deviation, and quartiles, allowing researchers an efficient and thorough view of quantitative data. However, educational psychologists have extended the power of calculators beyond the computing of statistics. Graphing calculators can support statistical research through plotting two-variable statistics; looking for patterns in two-variable situations to determine an underlying model to interpolate or extrapolate; and conducting various hypothesis-testing procedures, including *t*-tests, regressions, and analysis of variance (ANOVA).

To illustrate, let us suppose an instructor believes that attendance is an important factor in student motivation. A survey is given at the beginning of the course to determine beginning motivation levels (0–100 scale). The data are gathered, organized into

a list, and entered into the calculator. Various statistics can be quickly produced, including mean, standard deviation, and quartile boundary values. These numbers provide a clear picture of students' self-perception of motivation levels at the beginning of the course. In Figure 1, the middle column shows the mean (55.96) and standard deviation (14.87), whereas the right-hand column shows the five quartile values, including the minimum (20), median (58), and maximum (77).

```
EDIT CHLO TESTS    1-Var Stats        1-Var Stats
1: 1-Var Stats     x̄=55.962962296    ↑n=27
2: 2-Varr Stats    ∑x=1511            minX=20
3: Med-Med         ∑x²=90527          Q₁=42
4: LinReg(ax+b)    Sₓ=15.14921039     Med=58
5: QuadReg         σx=14.86602261     Q₃=68
6: CubicReg        ↓n=27              maxX=77
7↓ QuartReg
```
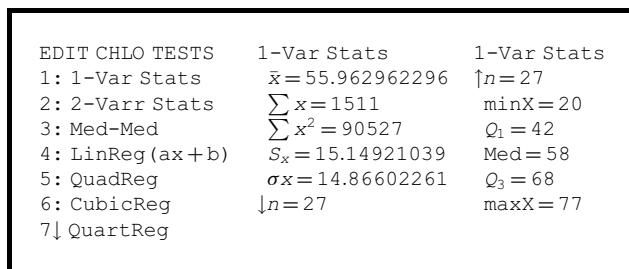
Figure 1   Students' Self-Perception of Motivation Levels at the Beginning of the Course

Although a four-function calculator could also produce these statistics, users would need to know the algorithms for computing the statistics, and input the quantities and operations in a lengthy, step-by-step process for each one. The built-in algorithms available in graphing calculators (or any calculator with these preprogrammed functions) allow users to produce these statistics quite easily.

Let us further suppose that, throughout the course, instructors implement the curriculum and record the students' attendance. Student motivation is measured again at the end of the course, and a new set of data is gathered and organized into another list. Changes in beginning and ending motivation levels could then be computed quickly for each student, creating a third list.

After entering the days present in class for each student into a fourth list, the instructor can then graph a scatterplot with days present as the independent variable and change in motivation as the dependent variable. With an appropriate viewing window, instructors can see a graph of the scatterplot (see Figure 2). Instructors can look at residuals to see if the model chosen is the most appropriate (see Figure 3).

Once the most appropriate model is determined, instructors can use it to predict attendance based on motivation levels. This can be very helpful in designing interventions for both student attendance and motivation.

In the past 10 years, there has been relatively little growth in the functional capacity of handheld calculators. Instead, a variety of communication, data storage, and photographic technologies have emerged as the choice of users of advanced handheld tools. Educational psychologists should look to the future to see how the functional capacities of current calculators will be incorporated into these emerging devices, and the degree to which these new devices will have an impact on the field.

David Slavit and Bill Kring

See also Descriptive Statistics; Inferential Statistics


```
LinReg                 Plot1  Plot2  Plot3
 y=ax+b                \Y₁=96630257959
 a=.9663025796         56X+-67.70415988
 b=-67.70415989        8448
 r²=.5928709054        \Y₂=
 r=.7699811072         \Y₃=
                       \Y₄=
                       \Y₅=
```
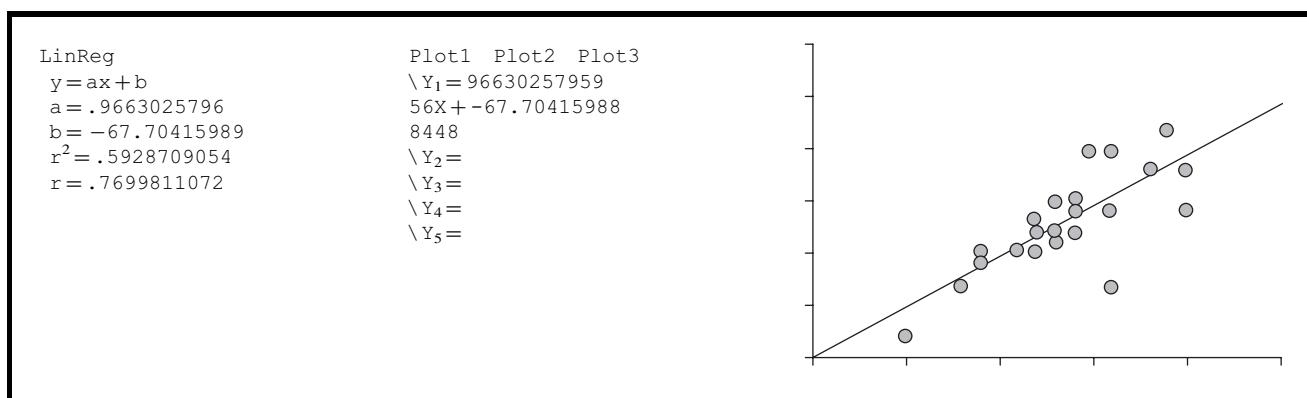
Figure 2   Scatterplot Graph
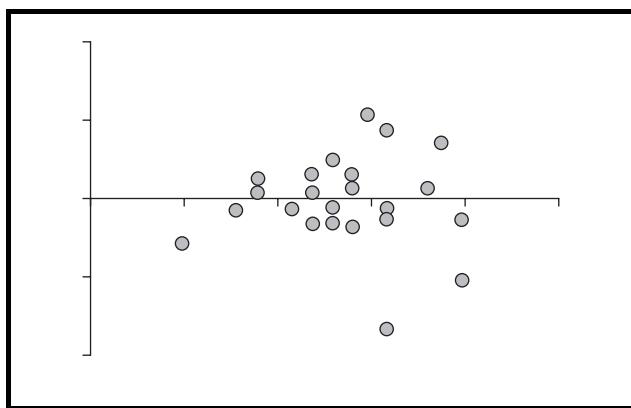
Figure 3   Residuals

Further Readings

Ball, G., & Flamm, B. (1997). *The complete collector's guide to pocket calculators.* Tustin, CA: Wilson/Barnett.

Kim, I. (1990). Handheld calculators: Functions at the fingertips. *Mechanical Engineering Magazine, 112*(1), 56–62.

Maxfield, C., & Brown, A. (2005). *The definitive guide to how computers do math.* San Francisco: Wiley.

Web Sites

Electronic Calculator History and Technology Museum: http://www.classiccmp.org/calcmuseum/calc.htm

Vintage Calculators Web Museum: http://www.vintagecalculators.com

CASE STUDIES

There are many different ways to study scientific phenomena, and one of the defining factors is whether the phenomenon, be it an individual (such as a child in the third grade) or an institution (such as a corporation), is studied in groups or individually.

The case study approach is a method used to study an individual, an institution, or any unique unit in a setting in as intense and as detailed a manner as possible. The word *unique* here is critical because the researcher is as interested in the existing conditions surrounding the object of study as much as the object of focus itself. In other words, if one were to study an 8-year-old child with a focus on her social development, one would, of course, spend a great deal of time on the child's psychological, emotional, and physical well-being, but also on her school relationships and the dynamics of her family. Similarly, one would not study the XYZ Widget Company by looking only at its profit-and-loss statement, but also looking at employee relations, policies, and other practices as well.

Many readers may have heard the term *case study* used before. The idea represents a major part of the methodology that physicians use to collect and disseminate information. The *Journal of the American Medical Association*, or *JAMA* (published weekly by the American Medical Association), regularly offers case studies of individuals whose conditions are so unusual that their symptoms and treatment demand special attention, and information about their cases needs to be disseminated. Also, the Harvard Business School regularly uses the case study model in its curriculum.

Physician-turned-psychologist Sigmund Freud pioneered the use of the case study in the development of his psychoanalytic theory. His famous patient, Anna O., and his detailed observations about her condition led to the use of free association as a method in the treatment of hysteria and other conditions. Also notable is the work of Jean Marc Itard, one of the first "special educators," and his case study description of the wild boy of Aveyron, which was the basis for the popular movie *The Wild Child.*

Advantages of the Case Study

Case studies offer several advantages over group studies of behavior, be it of an individual or an institution.

First, case studies focus on only one individual or one thing (for example, a person or a school district), and this allows close examination and the collection of a great deal of detailed data that can be further analyzed. Case studies are often used in a clinical teaching setting because of this advantage.

Second, because case studies encourage the use of several different techniques, a more varied collection of information becomes available. For example, personal observations, interviews, and other techniques all shed light on the focus of the study.

Third, there is simply no way to get a richer account of what is occurring than through a case study, which is what Freud did in his early work. He certainly could not have used a questionnaire to inquire about his patients' dreams, nor could he think to reach his level of analysis through the use of anything other than intensive scrutiny of the most seemingly minor details concerning the way the mind functions. These data helped contribute to his extraordinary insight into the functioning of the human mind and the first accepted stage theory of human development.

Fourth, case studies do not necessarily result in hypotheses being tested, but they do suggest directions for further study.

Disadvantages of the Case Study

Although a major tool in the qualitative methods kit, case studies do have shortcomings.

First, by their nature, they are limited in their generalizability. Studying one thing of anything does not easily apply to other, similar cases, be they individuals or institutions. The whole notion of a case study is to focus on the individual—by definition, indicating that individuals differ greatly from one another.

Second, although the case study model might appear to be easy to do (only one subject, one school, etc.), it turns out that this methodology is one of the most time-consuming research methods. Data have to be collected in a wide variety of settings and sources, and under a wide variety of conditions, and rarely is the choice available as to these settings and conditions.

Third, data may accurately reflect what is observed, but it is only one reality. Everyone comes to a given situation with a bias, and researchers must try not to let that bias interfere with the data collection and interpretation processes. The conclusions that are drawn based on case study science are based on a biased view of what is happening.

Fourth, what case studies provide in depth, they lose in breadth. Although they are extremely focused, they are not nearly as comprehensive as other research methods. As a result, case studies are appropriate only if one wants to complete an in-depth study of one type of phenomenon.

Finally, case studies do not at all lend themselves to establishing any cause-and-effect links between what is seen and the result. Although one can speculate, there is nothing in the case study approach that allows such conclusions to be reached.

Case studies do, however, reveal a diversity and richness of human behavior that is simply not accessible through any other method, but they have some serious shortcomings that have to be considered in investigations that scientists pursue.

Neil J. Salkind

See also Observational Learning; Piaget's Theory of Cognitive Development

Further Readings

Jennex, M. E. (Ed.). (2005). *Case studies in knowledge management.* Hershey, PA: Idea Group.

Manderson, L. (Ed.). (2003). *Teaching gender, teaching women's health: Case studies in medical and health science education.* New York: Hawthorne Medical Press.

CERTIFICATION

In the world of education, certification generally refers to the educational process by which future teachers earn the required state credentials to (a) teach specific subjects, (b) teach in specific areas, (c) teach specific grade levels, or (d) perform educational administrative functions in that state. This required teaching credential is often referred to as a teaching license or a teaching certificate depending on the term used by the department of education of the state granting the credential. Just as a driver's license can be categorized by the type of vehicle one is permitted to operate, so too, does a teaching license or teaching certificate specify what may be taught by the holder. Some certifications provide a license to teach specific subjects such as math, science, French, English, and so on, or to teach in specific grade levels or areas such as preschool, elementary (usually kindergarten to Grade 6), or special education, which can cover all grade levels and/or a specific need such as hearing impaired, gifted, and so on. There are also separate certificates for administration positions. Some of these are principal, curriculum supervisor, guidance counselor, or reading specialist. It is assumed that the holders of the teaching or administration certificates have met certain requirements that enable them to perform with knowledge and success in the specified domain.

Teacher Certification in the United States

In the United States, the responsibility for granting beginning teachers their licenses or certificates is borne by the state departments of education. Each of the 50 states has its own guidelines and qualification standards that must be met before a person is deemed certified to walk into a classroom to teach children. These state certification guidelines must be instituted by any university or college (private or public) that has a teacher education program within the specified state.

Because these guidelines are created independently by each state, there is often little consistency among the states' requirements. For example, in 2000, according to *Education Week*, 39 states required candidates for teacher certification to pass a basic skills test, but most of those states (36) had some sort of legal loophole that permitted some of the people who fail the exam to teach with a type of emergency or alternative certification. *Quality Counts 2000*, an annual 50-state report by *Education Week*, reported on teacher certification differences by stating that as a result of loopholes in state requirements, "millions of students sit down every day before instructors who do not meet the minimum requirements their states say they should have to teach in a public school" (p. 8).

An attempt to address this apparent lack of uniformity in qualifications for some teachers came with the passage of the No Child Left Behind (NCLB) Act of 2001, which was signed into law in January 2002. One portion of the 670-page NCLB Act of 2001 deals with "highly qualified teachers." This stipulation that each state must ensure that every public school elementary and secondary teacher is "highly qualified" before stepping in front of a classroom is a prime component driving certification standards. Highly qualified teacher requirements of NCLB are listed as a bachelor's degree, state certification, and demonstrated competency (as defined by the state) in each core subject that the teacher will be teaching. Many states are relying on standardized testing such as the PRAXIS™ series of tests to meet the "demonstrated competency" in each core subject. For many states, passing scores on various forms of the PRAXIS™ test in the core academic subjects that certification candidates intend to teach are required before certification is granted. Special education teachers and those who teach English-language learners must also demonstrate competency in the core subjects that they teach. (According to NCLB, these core subjects are English, reading or language arts, math, science, history, civics and government, geography, economics, the arts, and foreign language.)

Teacher preparation programs in colleges of education in the United States continue to vary from institution to institution despite the NCLB. Prior to NCLB, standards to improve the practice of teaching were created by various groups that include the National Board for Professional Teaching Standards (NBPTS), formed in 1987; the Interstate New Teachers Assessment and Support Consortium, known as INTASC, also formed in 1987; and the National Council for Accreditation of Teacher Education (NCATE) and the Teacher Education Accreditation Council (TEAC), which oversee teacher education programs throughout the United States.

One way in which certification standards are being raised for the beginning teacher is through initial licensure based on INTASC's 10 principles, which are focused on ensuring that teacher certification is a guarantee of a beginning teacher's knowledge and skill. These 10 principles are descriptions of a beginning teacher's competency in the areas of subject matter, learning styles, instructional methods, child development, learning environments, communication, planning, assessment, reflective practices, and community partnerships.

For experienced teachers, the obtaining of national board certification is a credential that affirms the expertise of accomplished teachers. National certification in teaching is awarded to expert teachers by the NBPTS and is designed to complement previously earned state teacher certification. The NBPTS, an independent organization with no government affiliations, lists five core propositions as the foundation of its standards: (1) commitment to students and learning, (2) knowledge of content and pedagogy, (3) management and monitoring of student learning, (4) reflection on the profession of teaching, and (5) participation in a learning community. The NBPTS offers 24 possible certificates in 15 subject areas and 7 age categories with two generalist certificates.

The rigorous requirements of the national board certification voluntary process were developed to promote and recognize excellence in teaching. In addition to registration fees, which are often borne by the local school districts of the applicants, candidates must provide an extensive portfolio (including videos of actual teaching) that documents accomplishments, excellence, and expertise. Candidates must also pass a series of essay-type tests as required by NBPTS. As of December 2006, there were approximately 50,000 teachers nationwide who were certified by the NBPTS.

Audrey M. Quinlan

See also Effective Teaching, Characteristics of; Expert Teachers; No Child Left Behind; PRAXIS™; Special Education

Further Readings

Berg, J. H., & Kelly, J. A. (2003). *Improving the quality of teaching through national board certification.* Norwood, MA: Christopher-Gordon.

Darling-Hammond, L., Wise, A. E., & Klein, S. P. (1999). *A license to teach: Raising standards for teaching.* San Francisco: Jossey-Bass.

Education Week. (2000). Who should teach? The states decide. *Quality Counts 2000, Editorial Projects in Education, 19*(18), 8–9. Available from http://www.edweek.org

Goldharber, D. D., & Brewer, D. J. (2000). Does teacher certification matter? High school teacher certification status and student achievement. *Educational Evaluation and Policy Analysis, 22*, 129–145.

Knutson, J. (1999). Teacher certification: Which end is up? *Educational HORIZONS, 79*(1), 2.

Raths, J. (1999). A consumer's guide to teacher standards. *Phi Delta Kappan, 81*(2), 136–142.

University of Kentucky College of Education. (2006). *Certification requirements for 50 states.* Retrieved December 28, 2006, from http://www.uky.edu/Education/TEP/usacert.html

U.S. Department of Education. (2004). Office of the Deputy Secretary. *No Child Left Behind: A toolkit for teachers.* Washington, DC: Author. Retrieved January 5, 2007, from http://www.ed.gov/teachers/nclbguide/toolkit_pg10.html#requirements

CHARTER SCHOOLS

Charter schools are public schools that are allowed greater autonomy than traditional public schools in exchange for increased accountability for meeting specific educational goals. Individual states' charter school laws vary tremendously, but charter schools generally operate as deregulated public schools, using public funds to support programs founded by parents, educators, community groups, or private organizations. State laws identify public entities like local school boards, universities, or state boards of education to evaluate proposals and grant a limited number of charters or contracts for establishing schools.

Charter schools are often launched to focus on a unique educational vision (e.g., Montessori), gain autonomy from embattled local districts, or serve a special population (e.g., children at risk of expulsion). The degree of autonomy enjoyed by charter schools varies but usually involves school-level decision-making authority over curriculum design, schedules, budget outlays, and hiring. The same entities that authorize charter schools are responsible for monitoring and ultimately closing charter schools that fail to demonstrate evidence of success by the end of a time period established by the charter, usually 3 to 5 years. Before discussing the potential and challenges of charter schools, it is useful to understand the origins of the charter school movement, the relationship of charter schools to the broader issue of school choice, and the popularity of charter schools.

The Charter School Movement

The charter school movement is young but has gained popularity as a mechanism for encouraging innovation and providing public school choice. The paragraphs that follow trace the growth of charter schools from their humble beginnings in Minnesota to their place today in mainstream public education.

Origins of the Charter School Concept

The earliest mention of the term *charter school* can be traced to the 1970s when a New England educator suggested that new educational approaches could be explored through contracts given to small groups of teachers. The idea was publicized in the late 1980s when a former president of the American Federation of Teachers suggested that local school boards could "charter" a school with union and teacher approval. Philadelphia dubbed their schools-within-schools initiative in the late 1980s "charter schools." In 1991, Minnesota passed the first charter school law developing a program to provide opportunity, choice, and responsibility for results. The following year, California followed suit.

Charter Schools and School Choice

The charter school concept has joined the ranks of school choice designs, including magnet schools, open enrollment, vouchers, and tax credits. Over the past 10 years, states and school districts have expanded opportunities for parents to use public funds to choose the schools their children attend in attempts to improve the quality of education available to students, particularly in urban areas.

Supporters contend that school choice can improve the public education available to all children based on two crucial factors. First, public school choice levels the playing field for less affluent families. Parents with financial resources can choose their children's schools through investing in private education or by virtue of the neighborhoods in which they choose to

live. School choice provides educational alternatives for families without these resources. Second, school choice proponents argue that competition among public schools serves the diverse needs of students more efficiently than a system requiring students to attend neighborhood schools. Market forces would, in theory, force all schools to improve in order to survive. When parents have the opportunity to choose strong schools, less effective schools lose students and ultimately close. These arguments have led to increasing popularity of school choice programs in general and charter schools in particular.

Criticism of School Choice

In spite of support for school choice from across the political spectrum, the concept is not without its detractors. Charter schools are a target of criticism because they are a prominent and growing example of the popularity of school choice programs. Detractors' concerns revolve around possible cultural isolation and racial resegregation in schools of choice as well as potential detrimental effects for students left behind in neighborhood schools.

A more balanced view suggests that school choice is not inherently good or bad, but its effects depend on policy decisions and their implementation. Big city school districts find themselves in a quandary for a way to gain consensus on a strategy for school reform. Charter schools are appealing because they appear to be a way out, but opponents' concerns should alert policymakers to potential drawbacks. Problems can be minimized through well-reasoned policy decisions in addressing the monumental challenges of improving the public education.

Popularity of Charter Schools

In spite of the controversy surrounding school choice, the popularity of charter schools has grown tremendously in a short period of time. Charter schools are a recent phenomenon, with the first school established just 15 years ago and the average age for all charter schools standing at only 5 years. Even so, they have quickly become a small but growing part of the mainstream educational system. In 2005, charter school laws were on the books in 40 states and the District of Columbia. Although laws varied dramatically from state to state, more than 3,000 charter schools were operating in the 2004–2005 school year.

Estimates of the number of children attending charter schools range from 750,000 to 1 million, or roughly 2% of all public school students. Nationally, the number of charter schools grew faster in 2004–2005 than in any of previous four years, adding almost 500 new schools. The states that opened the largest numbers of new charter schools were California, Colorado, Florida, Michigan, Minnesota, Ohio, and Wisconsin.

Great disparity exists among states in terms of the number of charter schools. With more than 500 charter schools in operation, California and Arizona top all other states in terms of sheer numbers. In fact, six states (Arizona, California, Florida, Michigan, Ohio, and Texas) account for almost two thirds of charter schools and charter school students. On the other hand, Washington, D.C. charter schools have the largest share of public school students of all states at almost one in four. In fact, although the total number of schools pales in comparison, a much higher portion of public school students in Washington, D.C., Delaware, and Colorado attend charter schools than in Texas or California.

Potential and Challenges of Charter Schools

The growth in charter schools comes from optimism that they can address many of the difficulties facing public education. Supporters anticipate that the autonomy afforded charter schools will foster innovation and ultimately lead to the development of better schools. Furthermore, they argue that increasing the number of schools from which parents can choose will lead to more balanced ethnic diversity across schools and improved educational outcomes. The hope and reality of each of these expected benefits of charter schools are explored in more detail in the paragraphs that follow.

Innovation in Charter Schools

Charter school supporters foresee innovation emerging from charter schools because of the autonomy they enjoy across many areas, including personnel, curricula, and schedules. Employees of charter schools do not typically belong to local teachers' unions and are considered at-will employees who can be fired for poor performance or compensated for superior service. In addition, alternative certifications may be an option for charter school teachers, bringing skilled individuals into

schools who may not otherwise consider teaching. Such an environment would be conducive to novel teaching methods, but the National Education Association argues that less stringent teaching credentials will lead to inferior educational quality. In addition, the association contends that charter school employees should have the same collective bargaining rights as any public school employees.

Flexibility to differ from district-mandated curricula contributes to the potential for innovation in charter schools. Many charter schools adopt specific curricular programs ranging from character education to technology themes. Freedom for charter schools to set their own schedules allows them to offer variations such as year-round schooling, extended days, or block schedules. In addition, many offer grade configurations unavailable in neighborhood schools, such as kindergarten through eighth or even twelfth grade. Charter schools are also more likely to use structures allowing student groups to remain with the same teacher for more than one year (called looping). Such unique programs are unlikely to be available in traditional public schools, giving parents real choice in finding the best schools to fit their children's needs.

Diversity in Charter Schools

As envisioned by school choice advocates, charter schools have the potential to create better school options for low-income and minority children in urban areas. Nationally, charter schools serve a higher proportion of minority and low-income students relative to traditional public schools primarily because more charter schools exist in urban areas. About a third of states require charter schools' ethnic composition to mirror that of their districts, and some require random processes like lotteries for admission when demand exceeds available seats. However, achieving racial targets is not a fundamental aspect of charter school operation as it is for magnet schools. In fact, some charter schools incorporate cultural heritage themes (e.g., African- or Native Hawaiian-based curricula), which results in less diversity within a given school.

Charter schools' success in providing viable alternatives for all parents is hindered because the families who would most benefit from the availability of public school alternatives often lack resources to take advantage of them. First, transportation or before- and after-school care can seem to be matters of convenience, but they are vitally important to some families' livelihoods.

Because few states provide transportation to charter schools, they are out of reach of those without a means of providing it themselves. As evidence of this, studies have shown that lower-income parents rate so-called convenience factors as more important, whereas higher-income parents placed more weight on a school sharing their philosophy.

Second, parents need information to understand and evaluate school options, and limited information access may create hurdles for the most disadvantaged students' families. Schools with a first-come first-served enrollment policy likely hinder families with less timely or sophisticated means of gathering information. As evidence of this dilemma, studies indicate that parents who participate in school choice have higher education levels than those who have a choice but fail to pursue it.

Finally, the reality of parental decision making can hamper efforts to increase racial diversity in schools. Studies suggest that parents report being most concerned about educational quality, but their behavior demonstrates a preference for schools that are close to home where their children are in the racial majority and the school body mirrors their own family's socioeconomic status. In fact, household race has been shown to be the strongest predictor of the racial composition of the charter school selected.

One largely unexplored aspect of diversity in charter schools is the degree to which they serve students with special needs. On average, charter schools enroll a somewhat smaller proportion of special education students nationally than traditional public schools. Once again, the data differ from state to state, with New Mexico and Ohio serving a larger proportion of students with special needs. This disparity stems from differences in goals established for charter schools in individual states. For example, some states emphasize charter schools serving students at risk of failure, which would tend to overrepresent students with special needs in charter schools. However, the concern exists that students who need additional services may be steered away from charter schools in a manner that would be discriminatory.

Educational Outcomes in Charter Schools

Charter school advocates claim that educational outcomes will improve as a result of creating alternatives to traditional public schools. However, the contradictory research findings on the academic success of

charter schools are staggering. Most studies on both sides of the debate suffer from serious methodological or data limitations. As with all educational research, identifying a reasonable comparison group is a major challenge. Two methods of creating comparison groups are reasonable but come with their own limitations. First, charter school attendees can be compared to students at comparable nearby schools. The problem is in controlling for the fundamental differences between parents who take advantage of choice when it is available and those who do not. Second, charter school attendees who won seats in a lottery can be compared to students who were not chosen. This design controls for the self-selection problem identified previously but limits analysis to schools with long waiting lists, which in itself introduces bias.

The tremendous variety of charter school laws across the country creates another fundamental challenge for researchers. Simply lumping together all charter schools regardless of their program or population and comparing them to the average public school is fraught with problems. Some charter schools target poor and disadvantaged students or students at risk of failure. Others seek to provide enrichment opportunities for students who excel. Some charter schools are funded comparably to nearby public schools, but most receive much less money. Finally, some charter schools operate in supportive local environments, whereas others must expend resources to defend their schools from antagonistic school districts or teachers' unions. Clearly, considering all charter schools as members of a homogeneous population is an oversimplification at best.

In spite of the difficulties, high-profile studies have been conducted attempting to answer questions about the educational success of charter schools. In 2004, the American Federation of Teachers and the National Assessment of Educational Progress reported that charter school students had lower achievement than public school students and that charter schools had a larger achievement gap between those who were and were not eligible for free and reduced lunch, a common proxy for family socioeconomic status. A subsequent Harvard University study encompassing 99% of charter school enrollments compared charter schools with schools students would most likely otherwise attend. This study found that charter school students were more likely than students in matched schools to be proficient in reading and math on state exams. Similarly, two national organizations with opposing views on charter schools reviewed existing charter school research and released contradictory findings in the spring of 2005.

As a result of the confusion in reconciling research on charter schools, the National Charter School Research Project is attempting to provide a balanced perspective and has committed to reviewing all available studies on the subject of charter school attendance and academic achievement. In all fairness, the state of research on charter schools is not surprising given the youth of the charter movement and the great disparity in programs across states. Definitive findings will come from an established body of work finding consistent results across settings and research designs. Such a body of knowledge is necessary to create a firm foundation for policy decisions, but it accumulates slowly over time.

Charter School Accountability

In exchange for increased autonomy, the charter school bargain requires increased accountability for results. However, differences in states' charter school laws lead to dramatic disparity across states in their respective charter schools' goals, standards of success, and consequences for failure. Generally speaking, charter schools can be closed if demand is low, if they fail to uphold their financial and operational commitments, or if they fail to satisfy the terms of their charter agreements. Even so, relatively few charter schools have been closed. During the 2004–2005 school year, only 65 charter schools, or 2% of the total, closed their doors in 17 states and the District of Columbia. As with all aspects of charter schools, the volume of closures varied from state to state, ranging from no closures in 15 states to 21 schools closed in California.

Both sides of the charter school debate attempt to use charter school closings to bolster their arguments. Supporters maintain that the small number of closures is a sign of charter schools' success, whereas opponents argue that few closures mean too many schools remain in operation regardless of performance. Whenever a charter school does close, supporters claim it is simply evidence that accountability works and consequences are real. Detractors contend that a closure is evidence of inherent flaws in the charter school concept. Ideology aside, closing a charter school is often difficult because even poorly performing charter schools can be tremendously popular with parents, especially when other available public schools are even worse.

Unfortunately, the autonomy enjoyed by charter schools allows room for misbehavior by inept and unscrupulous school leaders. Such was the case in the summer of 2004 when a for-profit, multisite Education Management Organization (EMO) in California faced closure of more than 60 campuses serving almost 10,000 students. In spite of overwhelming numbers, the California Charter School Association helped find new schools with little disruption for students. Students were distributed across California in small schools, and most were accommodated in one of the state's other charter schools. Although it is easy to place blame on the school management company, the local district contributed to the problem by failing to exercise oversight and ignoring ongoing poor performance.

To avoid public debacles like this, some states and the National Association of Charter School Authorizers are pushing pro-accountability standards for authorizing and monitoring charter schools. With tighter reins, intervention is possible before a problem grows to such a scale. For example, steps can be taken to strengthen a troubled school or make management changes at the first signs of problems. As counterintuitive as it may sound, some advocates suggest that the availability of more charter schools would reduce the impact of school closings because competition would draw students away from weak schools before a charter is revoked.

The Future of Charter Schools

In order for charter schools to realize their promise, advocates admit that many obstacles must be overcome. Charter schools have the potential to significantly increase the number of schools from which parents can choose, but a sufficient number of schools must be available to have an impact on the market. Achieving the scale necessary to successfully serve more than a small fraction of public school students requires changes to state laws, equitable funding, and plans for expanding proven models.

Charter School Growth

Even with their popularity to date, the future growth of charter schools is in question because of limits established in many states. Lawmakers often place limits on the growth of charter schools in their states as a means of political compromise. Twenty-seven states have restrictions on the growth of charter schools, most placing a ceiling on the number of new charter schools

that may open statewide, in cities, or under specific authorizing agencies. Other states limit growth by setting a maximum number of students enrolled in charter schools or limiting district spending on charter schools. Under constraints in place in 2005, only 725 more schools would be allowed across the country, and almost half of these (340) would be in California. Other leaders in the number of charter schools taken together, Michigan, Ohio, and Texas, are permitted just 29 additional schools under existing legislation.

Although legislative change will be required for the growth of charter schools to continue, pressure to lift existing limits may come from converting failing public schools to charter status. Specifically, provisions of the sweeping No Child Left Behind (NCLB) legislation allow low-performing schools to convert to charter status as part of their restructuring plans. If a school fails to achieve adequate yearly progress for 5 years in a row, it must be restructured in one of five ways: reopen as a charter school, replace the staff, contract with a private company for operations, allow state takeover, or implement some other major governance change. In the 2004–2005 school year, about 400 schools in 14 states have reached the 5-year mark, with another 750 schools in 31 states only 1 year away from reaching the limit. Even before NCLB would force restructuring, several districts, including Denver, New Orleans, San Diego, and Chicago, are beginning to convert failing schools to charter status.

The primary benefit of converting failing schools to charter status is the opportunity for new staff and a new program with the flexibility to address unique student needs. Not surprisingly, drawbacks exist to converting failing schools to charter status. In an attempt to find an easy solution to dealing with school failure, school districts could simply make the switch to charter status without making any substantive changes in the school operations. This strategy could buy time but would be unlikely to improve the school quality to benefit students. An honest commitment to restructuring a failing school as a charter school requires substantial district resources, from recruiting potential managers, administering the charter application and negotiation process, and involving the community to monitoring start-up and ongoing operations.

Funding Charter Schools

Funding inequities pose as serious a challenge to the future growth of charter schools as state caps.

Attempting to compare charter school funding to other public schools is fraught with many of the same issues as evaluating outcomes. For example, each state structures funding differently depending on student populations served and funding sources available. Furthermore, charter schools often pay for services not included in traditional public school budgets (transportation, oversight, etc.), and some districts provide services at no cost to district-run schools but not to charter schools (assessment, insurance, and services for students with special needs, etc.). Complicating matters, school district financial recordkeeping is not very sophisticated, so fiscal data are not readily available and accessible to researchers.

In spite of these issues, research that accounts for these methodological challenges makes it clear that the greatest funding inequity for charter schools comes from lack of access to incremental local budget dollars above state funds and from restrictions on funding for facilities. A 2005 study released by the Thomas B. Fordham Institute found charter schools underfunded relative to other district-run schools in the vast majority of communities examined, with the discrepancies being larger in most big urban school districts. The gap in 2002–2003 ranged from equity in Minnesota and New Mexico to gaps of more than 25% in Missouri, Wisconsin, Georgia, Ohio, California, and South Carolina. On a per-pupil basis, the average discrepancy is $1,801, which translates to almost a half-million-dollar shortfall every year for an average-sized charter school.

Achieving Scale in Charter Schools

In order for charter schools to become a serious force for improving education nationally, a sufficient number must exist to serve as viable competition for traditional public schools and to provide options for more than a fraction of students. Some charter school advocates laud the grassroots nature of many charter schools started by groups of committed teachers or parents. However, this approach alone is unlikely to produce the scale necessary to affect the quality of schooling nationwide.

One way of expanding the availability of proven charter school models is through EMOs. Most EMOs are for-profit companies like Edison Schools and Nobel Learning Communities, but most for-profit EMOs have yet to show a profit. Nonprofit EMOs like Aspire Public Schools exist as well. At present, only 10% of charter schools are operated by EMOs, and they tend to be larger than the average charter school. As with all aspects of charter schools, however, striking differences exist across states, with some restricting such organizations from operating in their territory by law. Alaska and Minnesota have no EMOs operating charter schools, whereas three quarters of Michigan charters are operated by EMOs, the largest proportion of any state. EMOs offer many resources for charter schools, including a leadership training ground, expertise and management systems, economies of scale, incentive and capacity to sustain schools over a period of time, and investments for research and development and possibly facilities.

Other designs attempt to merge the best of both worlds by leveraging the resources and support of experienced charter school organizers while maintaining the spirit of the independent charter school. Some examples include KIPP Academies in Houston and the Bronx, Minnesota New Country School/EdVisions, and High Tech High. KIPP Academies hope to help open 200 schools across the country by 2010. They are not part of an EMO because each school operates as an independent entity following the basic principles espoused by KIPP in its training and ongoing support programs. Minnesota New Country School/EdVisions is a high school design run by teacher cooperatives with learning through a personalized, project-based curriculum. The organization hopes to start 15 new secondary schools over 5 years with help from The Bill and Melinda Gates Foundation. High Tech High is a proven charter school model incorporating a rigorous personalized math, science, and technology curriculum with ties to the adult world. In 2006, The Bill and Melinda Gates Foundation awarded High Tech High $7.5 million over 3 years for new middle and high schools in California.

Replicating successful school models is likely to be the future direction for achieving charter school growth. Many envision fewer for-profit EMOs as well as a smaller proportion of individual start-up schools. Significant high-profile support exists for expanding the reach of charter schools. Organizations like The Bill and Melinda Gates Foundation are focusing on replication grants for expanding "successful" charter schools. The foundation's funding of charter schools is part of its ongoing commitment to fostering innovative educational programs and encouraging public engagement in improving education. The bulk (80%) of the support of the Walton Family Foundation,

which is devoted to education funding, is directed to charter schools. Their efforts range from direct support of new charter schools to funds for charter school management companies, technical assistance organizations, advocacy groups, and research. The founder of the Gap clothing store, Donald Fisher, is also an active contributor to charter schools. These influential philanthropists seem to share advocates' optimism in the potential for charter schools to improve education.

Future of Charter Schools

The charter school movement has gained popularity because of its potential to expand public school choice and improve public education. Predicted charter school benefits include innovation, higher-quality education, diversity, and accountability. The future of charter schools depends on policies that mitigate concerns about the possible downside of school choice. At the same time, further growth of charter schools requires individual states authorizing additional schools, equitable funding for charter schools relative to other district schools, and plans for achieving widespread growth of proven models. Although it is too early to declare charter schools a success or failure and intense study is just now beginning, charter schools' presence in mainstream public education and their high-profile support suggest that they are here to stay.

Angela K. Murray

See also No Child Left Behind; Virtual Schools; Vouchers

Further Readings

Betts, J., & Hill, P. T. (Eds.). (2006). *Key issues in studying charter schools and achievement: A review and suggestions for national guidelines* [Online]. Retrieved November 24, 2006, from http://www.ncsrp.org/cs/csr/print/csr_docs/pubs/achieve_wp.htm

Finn, C. E., & Hassel, B. C. (2005). *Charter school funding: Inequity's next frontier.* Washington, DC: Thomas B. Fordham Foundation.

Hamilton, L. S., & Guin, K. (2005). Understanding how families choose schools. In J. R. Betts & T. Loveless (Eds.), *Getting choice right: Ensuring equity and efficiency in educational policy* (pp. 40–60). Washington, DC: Brookings Institution Press.

Hassel, B. C. (2006). Charter schools: Mom and Pops or corporate design. In P. E. Peterson (Ed.), *Choice and competition in American education* (pp. 149–160). Lanham, MD: Rowman & Littlefield.

Lake, R. J., & Hill, P. T. (Eds.). (2005). *Hopes, fears, & reality: A balanced look at American charter schools in 2005* [Online]. Retrieved November 25, 2006, from http://www.ncsrp.org/cs/csr/print/csr_docs/pubs/hopes06.htm

Manno, B. V. (2006). Charter school politics. In P. E. Peterson (Ed.), *Choice and competition in American education* (pp. 149–160). Lanham, MD: Rowman & Littlefield.

CHEATING

Cheating involves an act of deception, fraud, or betrayal that often unfairly advantages the cheater over others. It can take many forms. The variety of behaviors it entails, and the wide range of contexts in which it occurs, is as diverse as the human species itself. From infidelity in the bedroom to malfeasance in the boardroom, people betray, trick, deceive, and defraud each other (and sometimes themselves) in a number of creative (as well as mundane) ways and places. This entry focuses on academic cheating and will describe several important facets or subcomponents related to the psychology of cheating.

Definitions and Typologies of Academic Cheating

A cursory review of the literature suggests that there is no universally embraced definition of academic cheating. For example, some researchers have defined cheating indirectly and vaguely, such as "a violation of an institution's policy on honesty," whereas others seem to have left the meaning of cheating up to students' interpretation by asking them directly how often they "cheat" on their work or use "cheat sheets" when they take tests. More typically, researchers have avoided such ambiguity or subjectivity, respectively, by asking students how often they have engaged in a specific set of behaviors, such as "copying from a neighbor during an examination" or "copying material without acknowledging the source." This latter approach is sometimes combined with a corresponding set of questions that asks students if they consider the behavior "cheating" or to rate how "serious" they think it is. Not surprisingly, the more likely students are to define a behavior as "cheating," the less likely they are to report engaging in that behavior.

In addition to the wide variation in how researchers have operationally defined cheating, several investigators

have created various typologies of cheating. Gary Pavela, for example, described four general types of academic dishonesty: (1) the use of unauthorized materials on any academic activity (e.g., using "cheat sheets" during an exam); (2) fabrication of information, references, or results (e.g., falsifying lab results); (3) plagiarism (e.g., copying verbatim another's work without proper attribution); and (4) helping others engage in academic dishonesty (e.g., allowing another to copy your homework). Stephen Newstead and his colleagues conducted an exploratory factor analysis on 21 academic behaviors and derived the following five factors: (1) plagiarism (which included a fabrication item); (2) collaborative cheating; (3) exams, collusion; (4) lying (e.g., lying about a medical condition to get an extension); and (5) exams, noncollaborative. More recently, some researchers have made a distinction between traditional or conventional cheating and digital or Internet-based cheating.

Academic cheating or dishonesty (the terms are often used interchangeably) has been defined in numerous ways, and various typologies have been constructed in an effort to map its vast terrain. Taken together, academic cheating can be defined broadly as the use of unauthorized or unacceptable means in any academic work. The means or actions include, but are not limited to, lying, using crib notes during exams, copying other people's work without permission, altering or forging documents, purchasing papers, plagiarism, unpermitted collaboration, altering research results, and providing false excuses to miss assignments or make up exams.

The Epidemic of Academic Dishonesty

Academic dishonesty is a pervasive problem in secondary and postsecondary institutions. By most accounts in the literature, the majority of students seem to be doing it and doing it in more than one way. For example, in its 2006 Report Card on the Ethics of American Youth, the Josephson Institute of Ethics found that 60% of secondary students reported cheating during a test at school within the past year (35% did so two or more times), and 33% reported plagiarizing material from the Internet (18% did so two or more times). Donald McCabe has found similar numbers in his national surveys of college undergraduates: 70% admit to engaging in some form of cheating, nearly 25% admit to "serious" test cheating,

and 40% to Internet plagiarism (77% of students don't believe such plagiarism is very serious). The problem is also prevalent at graduate schools, especially business schools, where 56% admitted to some form of academic dishonesty (i.e., copying other students' work, plagiarizing, or using prohibited materials on an exam) within the past year.

The high frequency of academic cheating is not a new problem—it's been labeled an epidemic numerous times since at least the 1980s—but there is evidence suggesting that it has grown over time. In "Schooling Without Learning: Thirty Years of Cheating in High School," Fred Schab documented the upward trend of academic dishonesty among high school students over the course of three decades. For example, in 1969, only 33.8% of students indicated that they had "used a cheat sheet on a test"; 67.8% admitted doing so in 1989. Letting other students "copy your work" moved from 58.3% in 1969 to 97.5% in 1989. Similar trends have been reported among college students. More recently, some have claimed that the Internet has *caused* a spike in plagiarism. Although such a concern is warranted (cutting and pasting is, after all, much quicker and easier than retyping word for word), research data suggest otherwise. First, plagiarism rates have not grown significantly, if at all, since the rise of the Internet in the early 1990s, and second, most students who report using the Internet to plagiarize also report using conventional means to do so. In short, the Internet does not appear to be creating a new generation of plagiarists and is probably best described as a conduit to, not a cause of, such cheating.

The Demography of the Dishonest

The title of this section, "The Demography of the Dishonest," is intentionally misleading. It reflects the common assumption that only certain "bad" people cheat, and the rest are "good" people who would never do so. The preceding section on the prevalence of cheating illustrates the faultiness of this assumption. The stark reality is that most students cheat at some point each and every year from middle school to graduate school. Therefore, the question isn't "Who does it?" (every type of student cheats—male, female; younger, older; Black, White; etc.), but rather "Who does it more often?" With this in mind, there are some demographic characteristics that have been associated with cheating.

In his meta-analysis of more than 100 published studies of undergraduate cheating (for the practical reason of access as well as the ethical requirement of obtaining parental consent, there are relatively few studies of cheating among middle and high school students; not enough, that is, to conduct meta-analyses and ascertain meaningful patterns), Bernard Whitley found age and marital status to be the most significant demographic predictors of cheating; specifically, younger and unmarried students were more likely to cheat than older and married students. Sex, parental financial support, on-campus residency, and number of hours of employment produced small effect sizes; specifically, cheating was more prevalent among students who were male, received more financial support from their parents, lived on campus, and were employed for fewer hours per week. Some researchers have used the terms *immaturity* and *lack of commitment* to describe or explain cheating among students with a combination of the demographic characteristics (e.g., young, unmarried, funded by parents, and unemployed).

Although academic ability or achievement, as measured by grade point average (GPA), is not necessarily a demographic variable, it has been included in numerous studies of cheating. As with most demographic variables, GPA seems to have only a small relationship to cheating in college; undergraduates with lower GPAs report higher levels of cheating. Again, data at the secondary level are scant, but according to the 29th Annual Survey of Who's Who Among American High School Students, 80% of the United States' best and brightest students reported cheating to get to the top of their class. This study and others suggest that the percentage of high-achieving students who cheat is comparable to that of lower-achieving students. It may be, however, that cheating among high-achieving students may be strategic and selective, and therefore "less" in an absolute sense.

In sum, there is no demographic profile for "cheaters"; most students cheat at some point. And although age, sex, marital status, and so on may be correlated with cheating, these demographic variables are almost never strong predictors of cheating. This is especially true when the conceptual model and statistical analyses of the study include psychological, social, and contextual factors. In other words, cheating is a complex problem. Any attempt to understand it more fully must go beyond demography and examine the more potent psychological processes that lead to cheating.

Morality, Motivation, and Misconduct

One of the most insidious aspects of academic cheating is that most students who report doing it also report believing that it is wrong to do. Why, then, do they do it? This question has been approached by numerous scholars in a wide variety of ways. The most direct way, of course, is to simply ask students why they cheat. Studies that have done so have yielded a fairly consistent pattern of results: pressure for grades, perceptions of poor teaching, time constraints, and lack of interest are typically among the primary reasons students cite when asked why they cheat in school. Although this approach may be the most straightforward and efficient, it is also atheoretical and superficial. For deeper insights into the question of why students cheat, even when they believe it is wrong, we need to turn to theory-driven studies that have employed more sophisticated research methods.

Some of the earliest and most prominent research on academic cheating focused on moral character to explain why some students did it and others did not. Contrary to expectations, Hugh Hartshorne and Mark May in their classic *Studies in Deceit* did not find evidence that honesty was a fixed, individual trait that reliably differentiated cheaters from noncheaters. Instead, situational factors, such as risk of detection and group approval, were the most influential factors in determining cheating behavior. Similarly, later studies that investigated the relations between moral development and academic cheating found that students with high reasoning ability cheated just as much as low-level reasoners when the threat of detection was low and the potential reward high. This is consistent with broader findings that moral judgment and moral action are not highly correlated, and the other components of moral functioning need to be accounted for.

Lawrence Kohlberg and his colleagues, for example, theorized that the relationship between moral development and moral action was mediated by two distinct but related types of moral judgment: deontic and responsibility. They described deontic judgment as a first-order judgment concerning the rightness or wrongness of a given action (deduced from a moral stage or principle), and responsibility judgment as a second-order affirmation of the will to act in terms of that judgment. Presumably, students who report cheating, despite having rendered a deontic judgment

that it was wrong to do so, did not make a judgment of responsibility; that is, they did not feel a sense of personal accountability to "follow through" and "perform the right action." Very few studies on academic cheating have directly tested this hypothesis, and those that have used the construct of "moral obligation" to do so. Conceptually, moral obligation and responsibility judgment are very similar, and, as expected, students who reported feeling a stronger moral obligation to refrain from cheating were less likely to report doing so.

Other psychologists as well as sociologists have also been interested in explaining the gap between moral judgment and moral action, and they have offered what might be considered the antithesis of Kohlberg's responsibility judgment or moral obligation: moral disengagement or neutralization of personal responsibility. Minimizing consequences (e.g., it's "no big deal"); euphemistic labeling (or nonlabeling, not acknowledging cheating as "cheating"); and displacing responsibility (blaming others) are three of the many disengagement mechanisms or neutralization techniques that individuals use to avoid or reduce self-recrimination when they have behaved criminally or immorally. Empirical research has demonstrated strong positive associations between cheating and moral neutralization. Displacement of responsibility seems to be the most prevalent strategy neutralization technique used by undergraduates: 61% of students who reported cheating rationalized their cheating by blaming others and/or some aspect of the situational context. Similarly, a study of high school students revealed that they were most likely to blame their cheating on teachers, and that this displacement of responsibility to the teacher was most pronounced among high-achieving and college-bound students.

Interview studies suggest that many students are aware of the incongruity between their beliefs about cheating ("It's wrong") and their behavior ("I do it"). They acknowledge the incongruity but are quick to dismiss it for the pay out of higher grades. In other words, students' academic motivation (pursuit of high grades) can trump their moral judgments. One of the most well-established approaches to understanding students' academic motivation is achievement goal theory, which posits the existence of two types of achievement goals: (1) to develop ability, often called a mastery or learning goal; and (2) to demonstrate ability (to avoid the demonstration of a lack of ability), often called a performance goal or ego goal.

Mastery goals orient individuals toward developing their knowledge, learning new skills, and using self-referenced evaluation criteria. Performance goals focus individuals on appearing smart relative to others, displaying skills, avoiding the appearance of inability, and using norm-referenced evaluation criteria. In educational settings, both types are posited to exist at three levels—school, classroom, and personal. Furthermore, students' goal orientations are thought to be situational, not dispositional—the extent to which a student is mastery and performance oriented is determined (at least in part) by the classroom goal structures, which are themselves partially determined by the school goal structures.

Over the past 10 years, numerous studies have used goal theory to further researchers' understanding of academic cheating. Studies at this intersection of goal theory and cheating have evolved over the past decade and fall into four basic types: (1) interindividual differences, (2) intraindividual differences, (3) longitudinal, and (4) experimental. In general, these studies have shown that mastery goals are negatively associated with cheating; students are less likely to cheat when they are focused on developing their competence and/or they perceive that classrooms or schools are focused on the development of their competence (these "messages" are communicated verbally through an emphasis on learning and effort as well as nonverbally through instructional and assessment practices). On the other hand, performance goals are generally positively associated with cheating; students are more likely to cheat when they, their classrooms, or their school are focused on demonstrating competence (through, for example, the attainment of high test scores and grades).

In sum, both morality and motivation matter in the perpetration of academic misconduct. In an interesting study that combined these typically distinct approaches to understanding cheating, Tamera Murdock and her colleagues used hypothetical vignettes to isolate the effects of classroom goal structures and teacher pedagogy on students' beliefs about the acceptability and likelihood of cheating. Consistent with previous research, they found that students believed cheating to be more justifiable (as well as more likely) when the classroom in their hypothetical vignettes was portrayed as focused on grades and the teacher as a poor instructor. They also assessed students' belief about the morality of cheating (i.e., an absolute, as opposed to context-dependent, belief

about rightness or wrongness of cheating). Not surprisingly, students' beliefs about the morality of cheating were less influenced by contextual factors and less strongly related to the perceived likelihood of cheating; students believe cheating to be morally wrong (regardless of circumstances), but this judgment doesn't much affect their likelihood of cheating.

The Power of the Situation

Individual behavior does not occur in a vacuum; social and situational circumstances exert a powerful influence on personal choices and actions. As discussed in the previous section, students' perceptions of classroom and school goal structure affect not only their cheating behavior but also their judgments about the acceptability of cheating. Also, as discussed above, students' perceptions of teachers' pedagogical competence affect cheating behavior and so, too, do their perceptions of teacher fairness and caring: Students are more likely to cheat when they perceive their teachers to be incompetent, unfair, or uncaring.

In addition to these subjective perceptions of teacher qualities, the subject matter also matters. Students report cheating most often in math and science courses and least often in social science and humanities courses. Susan Stodolsky's comparative analysis of math and social studies provides some insights into why students may cheat more often in math and science classrooms. Namely, the activity structures ("drill and kill") and assessment practices (objective, multiple-choice tests) often employed in teaching these domains provide more frequent and accessible opportunities to cheat. However, more research is needed. It may be, for example, that only certain types of cheating (e.g., homework and test-related behaviors) occur more frequently in math and science courses and that other types of cheating (e.g., plagiarism) are more likely to occur in social science and humanities courses.

Although all of the foregoing situational factors (real or perceived) have been significantly associated with cheating, peer norms (attitudinal and behavioral) tend to be the most powerful predictors of cheating behavior. For example, Donald McCabe and Linda Trevino's large-scale, multi-institutional study of individual and contextual factors associated with cheating in college found disapproval of cheating, peer cheating behavior, and fraternity/sorority membership to be the three most influential factors associated with cheating. Specifically, students who perceived that

their peers disapproved of academic dishonesty were less likely to cheat, whereas those who perceived higher levels of cheating among their peers and those who belonged to a fraternity or sorority were more likely to report cheating. Moreover, peers are very reluctant to report the cheating of others, even at institutions with so-called rat clauses that require students to do so. Put another way, cheating has become normative behavior among secondary and postsecondary students—it is widely seen and acceptable. Reporting others for cheating, in contrast, would be socially deviant behavior—rarely seen and greatly shunned. Finally, students who cheat rarely get caught. If caught, they are seldom punished severely, if at all.

Strategies for Promoting Academic Integrity

There are many ways that faculty, administrators, and institutions as a whole have attempted to address the problem of academic cheating in their classrooms or on their campuses. Many focus on the prevention and detection of cheating, but others have taken a somewhat different approach: the promotion of academic integrity. Although the promotion of academic integrity could be construed as a method of prevention, it is much more than that. Efforts to foster academic integrity, which are done most effectively at institutions with honor codes or committees, provide students with multiple opportunities and role models for learning the importance of understanding and concern for core academic values, such as honesty, trust, fairness, respect, and responsibility. Detecting and disciplining students who cheat must be part of any holistic approach, but it should not be the primary strategy to address the problem. Educational institutions at all levels should help students understand the meaning and importance of scholarship, intellectual property, and integrity. They should also help students develop the will and skill to participate in academic life in a fair, honest, and responsible manner.

Research indicates that students at institutions with honor codes are half as likely to report cheating on tests and one third as likely to report engaging in plagiarism. These reductions are far greater than any other approach known at this time. Unfortunately, relatively few institutions (particularly at the secondary level, where the problem of cheating begins to flourish and, therefore, the need is arguably the greatest) have invested the time and resources to transform the

culture of integrity on their campuses. This, despite the fact that there are now several organizations, such as The Center for Academic Integrity, that offer a wide range of materials and support for doing so. This being the case, the following strategies are offered as ways to mitigate the probability of cheating.

Reducing In-Class Test Cheating

The following are simple strategies that faculty can implement and should use, especially in larger lecture-based courses:

- Space Seating and Monitor: Where possible, students should be spread out (leaving a space between them), and faculty should move about the room throughout the examination.
- Create Multiple Forms: Whether or not spacing is possible, faculty should create multiple forms of their exams, randomizing both the order of questions and the answers.
- Ban Digital Technologies: Given the raise in digital forms of cheating, faculty should ban the use of cell phones, PDAs, calculators, laptops, and so on during exams.

Reducing Plagiarism

- Make Writing Assignments Clear and Manageable: Provide students with a list of specific topics or require components.
- Require Process Steps: Help students avoid last-minute, late-night plagiarism by requiring a series of process steps—topic identification, outline, first draft, peer exchange—that precede the final draft.
- Meet With Students to Discuss Their Research Papers: Where possible, meet with students individually to discuss their papers.

Detecting Plagiarism

- See the Signs: Make sure the paper addresses the specified topic or requirements, and take notice of changes in the voice or style of the writing, mixed citation styles or formatting, anomalies in diction, and so on.
- Know the Enemy: Faculty members should familiarize themselves with the online sources of plagiarism, such as Cheathouse, School Sucks, Screw School, and The Paper Store.
- Use a Plagiarism Detector: When plagiarism is suspected, faculty should use a text-matching program such as Turnitin.com to ascertain whether plagiarism has, in fact, occurred.

General Pedagogical Advice

Faculty can help students adopt mastery goals and develop academic efficacy with the following techniques:

- Engage: Create learning experiences that tap students' interest.
- Challenge: Provide optimal challenge and scaffold learning experiences.
- Empower: Give students a sense of control over the learning process and the products they create.

By taking the following actions, the faculty can help create an academic climate that emphasizes mastery over performance:

- Diversify: Design a multidimensional learning environment where expertise is distributed.
- Recognize: Emphasize and acknowledge students' efforts to learn and understand, not test scores.
- Privatize: Provide private individual evaluation of progress and avoid practices that invite social comparisons of performance differences.

In addition, the faculty can create a just and caring learning community by implementing the following actions:

- Play Fair: Establish and clearly communicate learning objectives and assessment practices; when possible, include students in curricular decision making.
- Care: Respect and support students in academic and nonacademic ways.

Jason M. Stephens

See also Peer Influences

Further Readings

Anderman, E. M., & Murdock, T. B. (Eds.). (2007). *Psychology of academic cheating.* Amsterdam: Academic Press.

Bandura, A. (1990). Selective activation and disengagement of moral control. *Journal of Social Issues*, *46*(1), 27–46.

Callahan, D. (2004). *The cheating culture: Why more Americans are doing wrong to get ahead.* New York: Harcourt.

Cizek, G. J. (1999). *Cheating on tests: How to do it, detect it, and prevent it.* Mahwah, NJ: Lawrence Erlbaum.

Jensen, L. A., Arnett, J. J., Feldman, S. S., & Cauffman, E. (2002). It's wrong, but everybody does it: Academic dishonesty among high school and college students. *Contemporary Educational Psychology*, *27*(2), 209–228.

Kohlberg, L., & Candee, D. (1984). The relationship of moral judgment to moral action. In W. M. Kurtines & J. L. Gewirtz (Eds.), *Morality, moral behavior, and moral development.* New York: Wiley.

McCabe, D. L., Trevino, L. K., & Butterfield, K. D. (2006). Academic dishonesty in graduate business programs: Prevalence, causes, and proposed action. *Academy of Management Learning and Executive*, *5*(3), 294–305.

Murdock, T. B., Miller, A., & Kohlhardt, J. (2004). Effects of classroom context variables on high school students' judgments of the acceptability and likelihood of cheating. *Journal of Educational Psychology*, *96*(4), 765–777.

Whitley, B. E., Jr. (1998). Factors associated with cheating among college students: A review. *Research in Higher Education*, *39*(3), 235–274.

CHILD ABUSE

Child abuse is a serious social issue that causes harm and results in the death of children every day throughout the United States and the world. It is both a psychological and an educational issue that affects individuals, families, schools, and communities, resulting in both immediate effects and long-term serious consequences on the lives of those it touches. This entry provides definitions, incidence of abuse, forms of abuse, characteristics of abusers, psychological effects on victims, and treatment interventions for both abusers and victimized children.

Definitions

Every state in the United States has a legal definition for child abuse and neglect based on the federal definitions cited in the Federal Child Abuse Prevention and Treatment Act of 1973 (CAPTA, 42 U.S.C.A. § 5106g) and as amended by the Keeping Children and Families Safe Act of 2003. Child abuse is defined as any act on the part of a parent, caregiver, or other individual that results in physical and/or psychological injury, or that can present a risk of serious harm to a child under the age of 18. Such acts include sexual abuse and sexual exploitation, and failure to act to protect the child if the abuse is witnessed. Neglect occurs when there is failure to provide for a child's basic needs for food, shelter, medical care, or appropriate supervision. Neglect includes failure to educate a child or attend to the child's educational needs, failure to provide psychological care, and knowingly permitting a child to use illegal substances (e.g., alcohol, illicit drugs).

Incidence of Child Abuse and Fatalities

The National Incidence Study of Child Abuse and Neglect (NIS) was mandated by the U.S. Congress to document and report on the incidence of child abuse and neglect. The most recent report, NIS-3 by Andrea Sedlack and Diane Broadhurst, covered a 7-year period and reported an increase in child abuse from previous years. Approximately 1.5 million U.S. children were abused or neglected during this period; physical abuse almost doubled and sexual abuse more than doubled. The most recent statistics available from the National Child Abuse and Neglect Data System (NCANDS) estimated 872,000 victims of child abuse in 2004; 1,490 of these were child fatalities (death of a child caused by injury from abuse or neglect); approximately 60% were victims of neglect, 19% suffered physical abuse, and 10% suffered sexual abuse. Children ages 4 or younger are most vulnerable because they are unable to defend themselves or seek help. Based on data from 32 states, 81% of the children who died were younger than 4, 11.5% were 4–7 years of age, 4.1% were 8–11 years of age, and 3.4% were 12–17 years of age.

The NIS-3 study identified factors that contributed to the highest incidences of abuse. These included the child's gender and age, family income, and family size. Race was not a significant factor found. There is little reporting of sexual abuse in children under the age of 3; however, the numbers begin to rise from preschool onward. Girls are more at risk for sexual abuse than boys; boys are more vulnerable to other forms of abuse that result in serious injury or death. Family characteristics contribute to higher incidences of child abuse. Children from single-parent homes have a higher risk of physical abuse and neglect than children living with intact families. Children who live in poverty or come from large families are more at risk of all types of child maltreatment.

Types of Child Abuse

Four major types of child abuse are recognized. These include physical, emotional, sexual, and neglect. These can occur separately, but often are found in combinations. Physical abuse results in a physical injury from acts such as severe beatings that can cause bruising, fractures, and death. Shaking, choking, throwing objects, kicking, burning, or other forms of physical acts constitute physical abuse. Punishment such as having children kneel for periods of time on stones or hitting them with belts or other objects is considered abusive.

Emotional abuse results from acts that include a continued pattern of degrading, denigrating, insulting, and ridiculing a child. It can include withholding affection and support, rejecting, terrorizing, and isolating the child from others. Some consider emotional abuse to be more detrimental to the child because it can have long-term psychological effects on self-esteem and psychological development.

Sexual abuse is sexual manipulation or coercion by an authority figure. It includes incest, rape, molestation, performing indecent sexual acts in front of children, exposing them to sexually explicit materials, photographing them in sexual explicit poses, and/or sexually exploiting them for prostitution.

Neglect is evidenced by the failure to provide adequate nourishment and by creating a living environment that is consistently dirty and unkempt. There is apathy on the part of the caregiver to provide for the child, and these children come to school disheveled with torn and soiled clothing. They are often left unsupervised, and their medical, psychological, and educational needs go unmet. Often, there is little or no affective stimulation or warmth, and in the case of infants, no nurturing or proper hygienic care is provided. Some victims of abuse have suffered multiple forms of abuse. For example, at times, sexual abuse may involve physical abuse, and a neglected child is most likely also emotionally abused.

Cultural Influences

Child abuse and neglect occur in all cultures and ethnicities, and on all socioeconomic levels. It is found in every country in the world. In some cultures, traditional values support the concept of physical punishment and the ridicule and humiliation of children for undesirable behaviors. What may constitute child abuse in the United States would not be considered so in other countries, especially in places where exploitation of children in the labor market, in war, or for sexual purposes is tolerated.

In the United States, African American, Pacific Islander, and American Indian or Alaska Native children have the highest rates of victimization; Asian children the lowest rates. According to the Child Maltreatment Report of 2004, White children were the most victimized at 53.8%; 25.2% were African American and 17% were Hispanic. Children with disabilities accounted for 7.3% of all victims. Parenting children with disabilities can increase the stress level for parents and caregivers, potentially placing these children at higher risks. Parental stress increases because these children can be more difficult to care for, have behavioral problems, have limited communication or mobility, and require constant attention and supervision.

Child Abuse in Schools

Child abuse is most associated with dysfunctional families; however, it does exist in our nation's schools. Some children experience their first incidence of abuse by teachers in schools. When abuse is observed in schools, it is sometimes left unreported by teachers and principals. Children often will not report experiences of abuse to teachers or other school professionals because they believe that nothing will be done about it. When they do report, they will disclose to peers or family members. It is critical for educators and school administrators to report disclosure of abuse by children and for school districts to have written policies on reporting procedures. In most states, educators who fail to report incidences or suspicions of abuse are subject to fines or other legal consequences, so professional development training for teachers and other school personnel is essential to help them understand signs and symptoms of abuse, reporting guidelines, and interventions.

The Abusers

There is no specific profile of an abuser, although various characteristics have been identified over time. Perpetrators of child abuse are defined by most states as parents or other caregivers such as relatives, babysitters, and foster parents. According to NCANDS's most current report, *Child Maltreatment 2004*, of the

872,000 child abuse victims, parents were the abusers in 78.5% of the cases; 6.5% were relatives; 4.1% were unmarried partners of parents; and the remaining were residential facility staff, child care providers, legal guardians, and foster parents. More than half of these perpetrators neglected their children, 10% physically abused them, and 6.9% sexually abused them. Fifteen percent committed more than one type of abuse.

Often, the perpetrators are young adults in their mid-20s living below the poverty level and lacking a high school education. They have difficulty coping with the stresses of daily living, and in some instances were abused themselves as children or were exposed to family violence. Family environments where these factors are present are more likely to expose children to frequent arguments, parental fighting (sometimes physical), and situations where children live in conflict and fear of violent reactions from their caretakers. Furthermore, parents and caregivers who themselves suffer from depression, personality disorders, substance abuse, or other emotional disturbances are more likely to abuse their children. Teenage parents who lack resources such as financial and social supports to care for children may also be at increased risk of abuse. Parents who are prone to act violently and who lack empathy perceive the world differently and may be more predisposed than others to be abusive to children.

Fathers tend to be the cause of the majority of child abuse fatalities resulting from physical violence, whereas mothers were most often responsible for neglect. Although parents tend to be the main perpetrators of aggression toward children, siblings and relatives living in the family environment can also physically, emotionally, and sexually abuse them.

Psychological Effects on Victims

Abuse is about control and power over another person. The psychological effects of abuse on victims will vary depending on the type of abuse endured, the length of time the abuse occurred, and the intensity of the abuse. A single abusive event can create Type I trauma, whereas repeated and prolonged abuse in children who endure severe pain, torture, consistent violence, and family conflict experience Type II trauma. Type II trauma has more debilitating effects on children. Children who experience Type II trauma develop coping mechanisms such as dissociation, denial, and shifts in mood from anger and rage to extreme passivity and withdrawal. Educators and human service professionals, when observing abused children, must consider the above factors and the extent of the trauma to better understand and treat symptoms and behaviors. Not all children will act the same. Some children will internalize the trauma and can be observed withdrawing and appearing sad, depressed, and lethargic. They may be fearful of others and avoid interactions. They may engage in somatization and exhibit physical symptoms such as headaches, stomachaches, or other ailments. They may exhibit self-mutilating and suicidal behaviors. Those who externalize the trauma will express their anger and rage toward others or animals. For example, they are prone to violence and hostility, and they may kill or torture animals or destroy property. In cases of sexual abuse, these children may demonstrate sexually provocative and explicit behaviors.

Effects of Physical Abuse

Physically abused children lack a joy for life. Depending on the extent of the abuse, they may have been removed from their homes and put in foster care or have endured multiple hospitalizations. In cases of severe physical abuse, these children may have physical impairments in speech, language, and motor functioning. They feel abandoned and rejected by their families. In school, they can exhibit academic and learning difficulties and rebelliousness, and have problems easily establishing friendships with other children. There may be physical consequences, such as bruises, burns, scarring, and broken bones. Abused children may suffer from depression and withdrawal, or may physically act out toward authority figures and peers.

Effects of Emotional Abuse

Emotionally abused children experience loss of self-esteem and often feel isolated and disengaged from others. They may become hypervigilant and suspicious and appear to be always on guard. In school, they may act belligerent toward teachers and may engage in hitting and fighting with other children. They are less motivated to attend school and complete academic assignments. Those who internalize the trauma will not engage in play or social activities and appear passive and nonresponsive to external stimuli. They can be moody and fearful of expressing feelings.

Effects of Sexual Abuse

Children who have been sexually victimized will exhibit an array of symptoms and behaviors that can run on a continuum from few symptoms to engaging in prostitution or becoming sexually predatory toward others. The behavior of these children depends on the severity of the abuse, the relationship to the perpetrator, the frequency and intensity, as well as the support from family members following disclosure. Common symptoms include depression, low self-esteem, anxiety, substance abuse, posttraumatic stress disorder, and self-destructive behaviors such as suicide and prostitution. These victims feel shame, embarrassment, and guilt. In extreme cases of repeated, prolonged sexual abuse, dissociation occurs as victims experience a sense of hopelessness and helplessness. When dissociating, the child copes by detaching from the self and the act that is being committed. A loss of identity and memory difficulties can result. These victims, because of their exposure to manipulated sex, become sexualized and will display overt sexually provocative behaviors, abnormal preoccupation with sex, and a lack of inhibition toward sexual acts or sexual conversation. Child abuse victims are at a higher risk for sexually transmitted diseases, pregnancy, and prostitution. They will have continued difficulty engaging in normal intimate and sexually rewarding interactions in adulthood.

Effects of Neglect

Victims of neglect have experienced a lack of, or no, nurturance and attention from parents or significant others. These children will suffer many of the same psychological effects as those who have been physically or emotionally abused. They will experience loss of self-esteem, abandonment, lack of affect, and difficulty expressing emotions. They tend to have a negative view of the world and may suffer from anxiety, depression, or overly aggressive behaviors toward others. They are often cognitively developmentally delayed. They feel no one cares for, and about, them and also experience a sense of hopelessness and helplessness to seek resources on their own.

Child Abuse Reporting

All states have designated agencies for the reporting and investigation of child abuse incidences, which then report these data to a central registry for tracking purposes. Specific state statutes require mandatory reporting of suspected child abuse. These statutes require certain professionals who work with children, such as educators, social workers, therapists, physicians, nurses, and law enforcement personnel, to be mandated reporters of suspected child abuse and neglect. Failure to adhere to these statutes can result in fines and legal charges, including felony charges and potential jail time. These professionals are believed to be in a unique position to observe instances of child maltreatment that might otherwise go undetected. Should a report be necessary, these professionals should inform the nonabusive parent that a report will be made.

If a minor child (client) discloses abuse, he or she should also be informed that a report will be made for his or her protection. Helping the victim deal with social service agencies and other authority figures who will be investigating the case and discussing the process is recommended. Early detection and reporting of maltreatment are critical in reducing further harmful effects to children.

Prevention and Treatment

The first step in prevention is reporting suspected abuse so it can be investigated and a determination made as to the most appropriate services for a child victim and his or her family. Referral to appropriate service agencies following the report provides support and resources to the abused child and family. Treatment should be comprehensive in that many resources and treatment modalities for all family members are provided. It is not enough to treat the abused child; family members must also be included because child abuse and neglect are experienced by the entire family.

Prior to selecting the treatment approaches that will be used, an assessment of the family is crucial. It should be determined whether parents or caretakers themselves have had personal experience with maltreatment in their family of origins, and if they may be suffering from personality disorders or other emotional disturbances that would warrant psychiatric intervention. It is essential to gain an understanding of parental child-rearing skills and stressful situations and events, and how these have been dealt with in the past within the family structure. An assessment of the child is useful to (a) determine if there is a disabling condition that could add undue stress to the family, (b) identify any behavior problems the child may have

at home or in school, (c) determine the child's developmental level, and (d) identify the child's adaptive and coping abilities. Knowing the child-rearing values and traditions of the culture of the family is important so as not to misinterpret social cues and behaviors on the part of the child, caregivers, and other family members. Observing the interactions between the child and caregivers can provide insight into the child–parent relationship and response behaviors toward the child. Parents and caregivers who maltreat children will exhibit less positive and nurturing behaviors toward them. They will be less likely to interact with their children, may be prone to quickly admonish or criticize the child, may have higher expectations than nonabusive parents, and may make inappropriate demands on their children for their developmental level. Following assessment, when selecting treatment approaches to use, therapists must make decisions regarding whether to treat the entire family as a whole, the couple, the abused child, or the parent–child dyad, and what specific modalities will be used.

Treatment Interventions for the Abusers

A behavioral approach has been recommended for abusive parents and caregivers because this approach is more directive, problem and solution focused, and educational in nature, and it places the responsibility on the family members to take action and make changes in their behavior. This approach is also intended to help parents become more sensitive and responsive to their children's needs. Cognitive therapies have been used to help parents gain awareness of irrational and dysfunctional beliefs and to change these belief systems into more rational, logical thinking that will promote appropriate behavioral responses. Individual, family, and group counseling can be used simultaneously. Individual counseling for the child and individual family members can be conducted in conjunction with group and family counseling interventions. Group counseling is beneficial because it connects parents to other maltreating parents who can serve as support networks to one another. In groups, parents can learn appropriate child-rearing and parenting skills and where to obtain assistance and resources for their families. Parents who maltreat children will need to learn techniques to control and manage their anger on a daily basis. They will learn how to respond to their children without shouting, hitting, or using other aggressive forms of behavior. They should be taught methods they can use to manage

and handle stress such as deep breathing and relaxation techniques. These individuals must be provided with opportunities to discuss feelings of frustration, shame, guilt, and inadequacy with their therapists and with other parents who share similar concerns. Role-playing by the therapists and group members can help ease parental anxiety and provide a mechanism by which to practice desirable behaviors and obtain immediate feedback from others. The therapist may need to combine treatment modalities with periodic home visits to assess whether what is being learned in therapy is being transferred to the home environment. To help abusive parents become more self-reliant, connecting abusive parents with sources of social support is necessary. These support venues include community agencies that can assist with housing, health, and social needs; churches or other religious organizations; day care providers; and employment counselors who can provide parents with resources to find better jobs or obtain training. Extended family members and friends can help provide emotional support and encouragement.

Treatment Interventions for the Victims

Treatment approaches for child victims will depend on the severity of the abuse the child has experienced. In addition, in order to be effective, the types of treatment selected must be congruent with the child's developmental level and symptoms. To protect and provide a safe environment, children are sometimes removed from the home and placed temporarily with relatives or in foster care. Therapeutic day care treatment has been used to increase the child's developmental and social skills and as a protective measure. The rationale behind the use of these two approaches has been to provide a hiatus for the family so they can participate in therapy before the child is returned home or adopted by another family. The intent has always been for this approach to be short term, but fewer than 40% of foster children ever return after 2 years in foster care. Furthermore, it is becoming increasingly difficult to find suitable foster care facilities for children, and many children shift from one home to the next and are separated from their siblings. Many abusive parents are never totally rehabilitated, nor do they respond to treatment positively, and they continue abusive patterns of behavior.

Child victims who experience severe trauma and suffer from symptoms of posttraumatic stress disorder are treated with education that involves teaching them

how to cope with stressful situations. Play therapy, the use of expressive arts and writings, and person-centered approaches that provide a means by which to express repressed feelings have been used. Experienced clinicians, in controlled settings, can also help the child recall memories of the traumatic event and work through feelings of anger, hurt, anxiety, shame, and distrust. This is especially critical for sexually abused children who have repressed memories of sexual exploitation and abuse.

Future Considerations

Child abuse is a social, educational, and psychological problem that occurs at epidemic proportions in the United States. It is a problem that warrants more attention than it has received in the past, both in the United States and abroad, because of the impact such abuse has in childhood and adult years. Education is needed to assist helpers and human service professionals identify signs and symptoms of abuse and to follow legally mandated reporting procedures. Treatment interventions aimed at helping both the abusers and victims effectively cope with the psychological factors resulting from abusive patterns of behavior and victimization are needed. Preventive measures, such as parenting skills classes and stress and anger management techniques, must be provided on a continuing basis by human service providers so as to educate others and to prevent abuse before it occurs. Efforts at prevention of child maltreatment are necessary in order for this epidemic to be thwarted.

Adriana G. McEachern

See also Conflict; Culture; Parenting Styles; Poverty; School Violence and Disruption

Further Readings

Abraham, N., Casey, K., & Daro, D. (1992). Teachers' knowledge, attitudes, and beliefs about child abuse and prevention. *Child Abuse and Neglect, 16,* 229–238.

Azar, S. T., & Wolfe, D. A. (2006). Child physical abuse and neglect. In E. J. Mash & R. A. Barkley (Eds.), *Treatment of childhood disorders* (3rd ed., pp. 595–646). New York: Guilford.

Child Welfare Information Gateway. Child Abuse Reporting Numbers. Retrieved September 7, 2007, from http://www.childwelfare.gov/pubs/reslist/ rl_dsp.cfm?rs_id=5&rate_chno=11–11172

Howe, D. (2005). *Child abuse and neglect: Attachment, development, and intervention.* New York: Palgrave Macmillan.

Kenny, M. C., & McEachern, A. G. (2002). Reporting suspected child abuse: A pilot comparison of middle and high school counselors and principals. *Journal of Child Sexual Abuse, 11,* 59–75.

Pritchard, C. (2004). *The child abusers: Research and controversy.* New York: Open University Press.

Runyon, M. K., Kenny, M. C., Berry, E. J., Deblinger, E., & Brown, E. J. (2006). Etiology and surveillance in child maltreatment. In J. R. Lutzker (Ed.), *Preventing violence: Research and evidence-based intervention strategies* (pp. 23–47). Washington, DC: American Psychological Association.

Schwartz, L. L., & Isser, N. K. (2007). *Child homicide: Parents who kill.* Boca Raton, FL: Taylor & Francis.

Sedlak, A., & Broadhurst, D. (1996). *Third national incidence study of child abuse and neglect: Final report.* Washington, DC: U.S. Department of Health and Human Services, Administration of Children and Family.

Skinner, J. (2001). Teachers who abuse: The impact on school communities. *Educational Research, 43,* 161–174.

Wolfe, V. V. (2006). Child sexual abuse. In E. J. Mash & R. A. Barkley (Eds.), *Treatment of childhood disorders* (3rd ed., pp. 647–727). New York: Guilford.

Web Sites

Child Welfare Information Gateway: http://www.childwelfare.gov

U.S. Department of Health and Human Services, Administration for Children and Families: http://www.acf.hhs.gov

CLASSICAL CONDITIONING

Classical conditioning (also named *Pavlovian conditioning*, for Ivan Pavlov, the researcher who pioneered the topic) is considered, along with habituation, to be a fundamental form of learning. The term *conditioning* is used because classical, along with instrumental and operant, is considered a simpler and more basic form of learning. As a form of learning, classical conditioning involves a relatively permanent change in behavior potentiality that arises as a result of particular experiences. Hence, at one time a certain behavior does not take place. But as a result of independent variables (or imposed stimulus contingencies), a particular target behavior then does take place with greater probability.

Unlike habituation, which involves only a single stimulus, classical conditioning involves two stimuli that are paired. The purpose of one stimulus is to reliably elicit behavior. This is the *unconditional stimulus* (or US, because it will provoke a response unconditionally), also called *unconditioned stimulus* (because the response to it did not have to be trained). It elicits the unconditional, or unconditioned, response (UR), which is already in the organism's repertoire. The other stimulus—the *conditional stimulus* (CS), because the response to it is conditional, or *conditioned stimulus*, because the response to it must be learned—is neutral with respect to the US. The CS, in order to be a stimulus, must evoke some response; but that response, or pattern of responses, is unrelated to the US, and hence, those responses have little or no similarity to the UR. If training is effective, then the CS has acquired some capacity to elicit a response that resembles the UR. But because it is elicited by the CS alone, it is considered a learned response—a conditioned (learned), or conditional, response (CR).

Training consists of pairing CS and US, ideally with CS onset preceding US onset on each trial. So, initially, before conditioning, the CS provokes little or no conditioned response, which serves as a baseline for evaluating the amount of learning reflected in the amount (percentage, magnitude, probability) of CR following training.

All of the basic phenomena that take place in learning paradigms (to include instrumental, operant, and habituation) are found in classical conditioning. As in other paradigms, acquisition occurs as a negatively accelerated function, meaning that the greatest behavior change takes place during early trials. A practical implication is that those initial trials are rather crucial because this is the time when errors are readily mastered, and considerable training will be needed to extinguish those flawed behaviors. Another implication arises later in learning when performance is nearing asymptote. As a curve of decreasing returns, less and less additional change in performance (reflecting less and less further learning) is observed with additional training trials. (Good examples of these considerations are found in learning physical or sport skills. During initial training, mistakes can be learned quite readily—pointing out the importance of guidance and error-free training for novices. Later, highly skilled performers have difficulty maintaining motivation to keep working when

their efforts seem to show very little payoff.) The other basic phenomena include such matters as extinction, spontaneous recovery, stimulus generalization (positive transfer), stimulus discrimination, external inhibition, disinhibition, and so on.

A great deal of research has been conducted dealing with time relations between CS and US presentations and characteristics of stimuli (e.g., intensity and complexity). It is quite clear that for optimal conditioning, the CS onset should precede US onset—the time between onsets being commensurate with the latency of the UR. For a very simple and quick response (e.g., eyeblink, which is popular in much of the animal research), roughly 0.5 seconds seems optimal, whereas for a slower or more elaborate response, some number of seconds would be suitable. In essence, the CS is functioning as a signal that forecasts the presentation of the US. So, it appears as though there needs to be sufficient time for the CS to allow the learner to prepare for the US, and to anticipate it. When it comes to stimulus intensity, it appears that US intensity is of particular importance. It must be "prepotent" in that it is strong enough to take priority, dominating over other stimuli that could cause distraction or evoke behavior that could compete with and weaken the UR. The function of the US is to evoke a particular pattern of behavior, so it makes sense that a strong and unambiguous elicitor is desirable. A weak US will evoke a weak UR or a UR that could be viewed as "sloppy" in the same way that unclear and ambiguous guidance or instruction leads to behavior that is not clearly on target. With classical conditioning, the behavioral outcome is thoroughly in keeping with the behavior evoked with the US. Or, in other words, "what you do is what you learn." It should be noted that a US need not be appetitive (evoking behavior that fits a general category of "approach") but can be aversive (leading effectively to avoidance behavior).

Higher intensity may be of some importance for the CS—but within reason. An excessively strong signal can provoke some behaviors that can interfere with both CR and UR. It has been suggested that any stimulus above threshold (detectable) can serve effectively as a CS (as long as it reliably predicts occurrence of a US)—meaning that learning can take place with very subtle cues. But to have an optimally effective CS, it would appear that not just adequate intensity is important but greater complexity is of considerable value. Note that complexity does not mean difficulty; rather,

it means number of elements. So, when signals are important, they include many components (e.g., a railroad crossing in a high-traffic area or warnings about fuel and altitude in aircraft cockpits will have a variety of visual features plus sound). Typically with compound stimuli, a particular element will be dominant; sometimes, that element will overshadow the others. But typically, the "weaker" elements will add and contribute to the overall signal value.

Classical conditioning can easily be regarded as a substrate in that it can be seen as a basis or underlying process in instrumental and operant conditioning. It also has rather obvious roles in those paradigms whenever some learned element is involved. For example, secondary reinforcers (also appropriately called conditioned reinforcers) are extremely powerful and practical tools for instrumental and operant conditioning. Perhaps the ultimate example is clicker training for dogs and other animals, where the sound of the clicker (which serves as the secondary reinforcer) had been paired with some incentive (dog treat), but can now serve to "reward" the animal immediately following some target behavior. Virtually all of the social reinforcers for humans (smiling, nodding, attention) can be seen as having acquired their incentive value in analogous fashion. Discriminative stimuli in instrumental and operant conditioning serve as signals that have been associated with particular contingencies rather than specific stimuli.

Finally, in practical application, classical conditioning is often where some form of training begins as a way to establish some behavior. And once some behavior is stable and reliable, it can be used for building additional outcomes. For example, with higher-order conditioning, an established CS can function as a US. Furthermore, it probably rarely maintains a "pure" form but becomes combined with or replaced by instrumental contingencies and incentives. For example, if a parent is teaching vocabulary, the parent may prompt by saying, "Say ball." After a few of these trials (US-UR), the parent may hold up a ball, point to it, and say, "What's this? Say 'ball.'" Then, the parent may fade out the US, eventually presenting the CS (pointing to the ball and saying, "What's this?"). When the child emits the CR of "ball," then the parent may respond with a social reinforcer: "That's right; good for you." It could be argued that classical conditioning should be seen as the substrate and underlying process for most of our learning. Indeed, most of our vocabulary and nearly all our emotions (preferences as well as fears) can be attributed to classical conditioning.

Louis G. Lippman

See also Behavior Modification; Learning; Learning Style

Further Readings

Tarpy, R. M. (1997). *Contemporary learning theory and research.* New York: McGraw-Hill.
Turkkan, J. S. (1989). Classical conditioning: The new hegemony. *Behavioral and Brain Sciences, 12,* 121–179.

CLIQUES

Relationships outside the household become increasingly important as a child moves through adolescence. One way to view these relationships is through cliques, which can be defined in two ways: (1) Researchers define a clique as a grouping of persons who interact with each other more regularly and intensely than others in the same setting (a more neutral definition), and (2) the more popular definition by laypersons is a social grouping of persons that exhibits a great deal of peer pressure on its members and is exclusive, based on superficial differences (a more negative definition).

No matter which definition is chosen, cliques exhibit common criteria. Cliques are a small social group of anywhere between 2 and 12 persons, typically averaging around 6. They are created based on similar characteristics of their members, including age, gender, race, social status, socioeconomic status (SES), and shared interests and activities. Cliques provide the major social context in which youths interact with each other, and thus are of high importance for their healthy development.

Along with the definition, an important distinction must be made to separate cliques from the similar peer grouping of crowds. Crowds are a larger grouping of individuals who share a common image or reputation with each other, or may have similar features (race, SES), but do not necessarily accept their membership in said group. Examples of crowds are jocks, nerds, populars, cheerleaders, and so on.

This entry describes the formation and maintenance of cliques, including gender differences, changes in clique structure across adolescence, types of individuals

in/out of cliques, the process of joining cliques, as well as the benefits and problems with cliques.

Gender Differences in Cliques

Although many similarities exist between male and female cliques, they do have some important differences. Females are more likely as a whole to become members of cliques, whereas boys are more likely to be nonmembers. Females are also more likely to form cliques at an earlier age than boys, with females beginning around age 11, boys at 13 or 14. Female cliques are more focused on vocal activities, including gossip and sharing thoughts and feelings, and are easily hurt by the gossip and hurtful remarks spread by these cliques. Male cliques are typically focused around similar activities and interests. Research has also found that males have a greater desire for acceptance from a group, and more concern with status than females, who were more interested in being emotionally closer to a smaller group of peers.

Changes in Clique Structure Over Time

It has been stated that the structure of cliques develops in parallel to the development of identity; thus, important changes are seen as youths move through adolescence into adulthood.

Preadolescence

Preadolescence marks the initiation of cliques as youths begin to separate from familial relationships being of primary importance. These early cliques are created almost entirely of same-sex individuals and are formed around similar interests and activities rather than other demographic characteristics.

Mid-Adolescence

Mid-adolescence involves great changes in both clique structure and activities, as youths become more interested in the opposite sex, but do not yet have dating relationships. This is a time in which cliques become more segregated based on demographics beyond sex and interests/common activities. Cliques begin to integrate at similar venues with opposite-sex cliques, but their membership does not yet change. Later, as youths become interested in dating relationships, cliques begin to open up to mixed-sex and even mixed-age membership.

Late Adolescence

As youths begin to enter late adolescence, dating relationships begin to take precedence over clique membership. Cliques begin to collapse as members are lost and are replaced with pairs of dating adolescents. These pairs form loosely associated cliques, which separate from the larger group activities in favor of individual ones. This form of clique structure is said to extend into adult relationships.

Types of Individuals in and out of Cliques

Through the use of observation studies, researchers have identified three types of youths in clique membership: clique members, liaisons, and isolates.

Clique Members

Clique members comprise fewer than half of all adolescents, and are actively a part of a single, small group of persons. They spend most, if not all, of their social interactions with these same individuals.

Liaisons

Liaisons are individuals who interact with some members of a clique, but not with all. Liaisons can associate with multiple separate cliques, and are generally well thought of by their peers. This type of individual comprises approximately 30% of youths.

Isolates

Isolates are individuals who are not involved with any clique members, and generally have few, if any, relationships with peers in their social network. Isolates may be of volunteer status, in that they actively avoid relationships, or they may be of forced status (also labeled as targets or victims), in which they are set up by others to be excluded and ridiculed.

Joining Cliques

Because of the hierarchical interior structure of cliques, some individuals who desire membership are denied. This membership is maintained through

a careful screening process, usually established and maintained by the leader. Two methods for obtaining membership exist: invitation and application.

Invitation

Invitation into a clique can occur as a direct request from one of its current members or by indirect socialization. Indirect socialization can resemble an advertisement of what the clique can offer the prospective member, involving a courtship attempting to show how great life in the clique actually is. Also important (with application as well) is that before this potential new member is invited/accepted, the leader(s) of the group must desire his or her membership; however, if the leader(s) of the clique want to give someone membership whom the other members disapprove of, the person may still receive an invitation/acceptance with little effort from the leader(s).

Application

Application is the second way a person can be granted entry into a clique, which may not appear much different from the outside looking in, where a similar courtship will exist. Also important is that although application may not be immediately made through the leader (applicants may begin to socialize with lower-level clique members), nearly all applicants will end up attempting to impress the leader(s) before acceptance is granted.

Status and Hierarchy of Cliques

Maintaining Status

Cliques have a very strong hierarchical status structure inside them, containing at least one leader (typically a single person reigns). This status structure can also be quite transient, and changes in status are likely to occur regularly. It is said that members need to be constantly working to hold on to, or increase, their position in the clique. For leaders of cliques, this can become a harsh process, which has led many to hold a negative view of cliques. One way in which a leader maintains his or her status is to alternate praise and criticism toward other members, which keeps the leader from gaining negative attention on him- or herself. Leaders may also change the way the clique views activities, values, and/or opinions of things to keep the other members unsure of where

they stand in the group. Beyond these changes, clique members constantly change friendships so as to socialize with only the most popular of peers.

Types of Group Members

In her book *Queen Bees & Wannabes: Helping Your Daughter Survive Cliques, Gossip, Boyfriends & Other Realities of Adolescence*, Rosalind Wiseman describes her observations of cliques and groups female adolescents into several roles:

Queen Bee: The leader of the group, who Wiseman describes as ruling by "charisma, force, money, looks, will, and manipulation" (p. 25).

Sidekick: The lieutenant of the clique, who is said to have no voice but that of the Queen Bee, consistently backing her up no matter what the issue.

Banker: The girl who holds all of the information and gossip, and releases it for her own benefit.

Floater: Similar to the liaison described earlier, a girl who can go in and out of more than one clique.

Pleaser: This person can be in or out of the clique; she will consistently back up the Queen Bee and the Sidekick, but receive very little credit for doing so.

Target: Typically outside the clique, this person is excluded by the members and humiliated on a regular basis.

Male Roles

Although Wiseman's book centers around the lives of female teenagers, she also adds a section for males, and describes their similar roles:

Leader: Similar to the Queen Bee, Wiseman describes him as, "Everyone wants to be him … athletic, tough, able to get the girls, or rich" (p. 183). Contrary to the Queen Bee, he is typically well respected.

Flunkie: Similar to the pleaser, in that he will do anything asked of him, but different in that he responds to any member. This person constantly gets into trouble and irritates others because of his actions.

Thug: Can also be a wannabe thug. This person is smarter than he lets on and typically communicates in nonverbal bullying behavior. Although this person appears popular, he may or may not be.

The Get Wits: The groupies of the male clique. These males are well respected by adults for being good, but

their peers do not hold them in similar respect. and they are seen as tagalongs.

Hierarchy of Cliques in Social Setting

Not only does a hierarchy persist inside of cliques, but a hierarchy also ranks cliques belonging to the same social network. Similar to the hierarchy and status maintenance that occurs inside of cliques, there are also changes to the clique rankings; however, this status rank does not change as frequently as internal statuses. Although cliques differ from crowds, these larger peer groups can play a large role in the status of a clique. The more popular the crowd with which members of the clique identify, the higher the status of the clique, and multiple cliques identifying with the same crowd will be more likely to have status fluctuations with these similar-status cliques.

Benefits and Problems With Cliques

Benefits

Although cliques are typically seen as a negative force on adolescents, they do have some positive attributes. One such attribute involves the development of social identity. Through the formation of peer groups such as cliques, adolescents learn to identify types and regulate social interaction. Adolescents begin to see themselves as others perceive them, and they are taught through these interactions to control their emotions and behaviors.

A second benefit of cliques is that they can help promote an adolescent's self-worth. Research has shown that they provide a strong effect on this important feature of self-esteem. As stated earlier, adolescents begin to see how others perceive them through peer groups, and the perception of others has been found to have a large impact on how a person views him- or herself.

Problems

The benefits with cliques can also easily become problems. With regard to the development of social interaction, adolescents in overly competitive cliques can perceive all social networks as being this way, and thus act in a way to compete with others. With self-worth, adolescents can easily develop negative perceptions of how other people perceive them, or they may see the constant changes in status and think

that they are worthy only if they have something to offer or if they push other people down.

For better or for worse, peer groups such as cliques are the most prevalent friendship structure in adolescence and thus produce a strong effect on adolescents' further development.

Matthew J. Davis

See also Bullying; Friendship; Identity Development; Peer Influences; Psychosocial Development; Social Development

Further Readings

Adler, P. A., & Adler, P. (1998). *Peer power: Preadolescent culture and identity.* New Brunswick, NJ: Rutgers University Press.

Moehn, H. (2000). *Everything you need to know about cliques.* New York: Rosen.

Rice, F. P., & Dolgan, K. G. (2004). *The adolescent: Development, relationships, and culture* (11th ed.). Boston: Allyn & Bacon.

Steinberg, L. (2005). *Adolescence* (7th ed.). New York: McGraw-Hill.

Wiseman, R. (2002). *Queen bees & wannabes: Helping your daughter survive cliques, gossip, boyfriends & other realities of adolescence.* New York: Crown.

COGNITIVE AND CULTURAL STYLES

The unique ways in which students learn and teachers teach, or learning and teaching styles, are closely related to cultural values and personality types. Matching preferred styles can prevent school failure, enhance success, and motivate students to stay in school and develop talents to their fullest. Matching students' preferred styles can also make the teaching experience more satisfying for educators and can increase the effectiveness of educational programs. However, as American society becomes more complex and diverse, schools and teachers are finding it more difficult to identify and match the preferred cultural and cognitive styles of students.

Cultural styles reflect virtues and philosophies of life that are emphasized by families, communities, and cultures. As one component of learning styles, they serve as guideposts or markers that children use as they move through life in search of the careers and the life goals they find meaningful and fulfilling.

A second major component of students' unique learning styles and teachers' instructional styles is cognitive styles. These are styles of personality that determine how students like to learn, the ways in which they prefer to relate, the types of rewards that make success in school meaningful, the preferred manner of communication, and leadership style. Cognitive styles are related to cultural styles through the process of socialization and instruction by parents, other authority figures, and cultural experiences in the home and community.

By the time children attend school, they have developed specific cultural and cognitive styles that are related to how they like to learn and how they process and retain information. These cultural and cognitive styles may or may not be compatible with their instructors' teaching styles and the cognitive and cultural styles emphasized by the schools they attend. Some cultures and families emphasize emotional IQ, or the importance of understanding people and relationships, and emphasize being a helpful and spiritual person who is a contributing member of a family and society. Instruction by adults is largely done by modeling and demonstration while simultaneously discouraging deviation from set practices and procedures. Other families and cultures tend to emphasize the value of traditional educational skills such as reading and math, and the importance of the individual is emphasized over that of the group. Instruction in these families is largely transmitted through trial-and-error learning where the child is encouraged to learn and to work independently of adults.

This entry focuses on how teachers, schools, colleges, and universities can become more sensitive to the preferred cultural and cognitive styles of students regardless of familial and cultural background. It is argued that the ultimate goal of education should be to identify the uniqueness of learners, and to individualize instruction in order to match the individuality of expression in learning environments. Specifically, the goal should be to match preferred student styles and worldviews in order to ensure enthusiasm for learning and success, thereby enhancing the adaptability and flexibility necessary to live happily and meaningfully in a diverse society and a global world.

Cultural Styles

Manuel Ramirez and Alfredo Castaneda's theory of cultural and cognitive styles flexibility proposes that cultures, communities, and families hold certain belief systems and perspectives on the meaning of life, identified as cultural styles, which can be classified on a traditionalism-modernism continuum. Bicultural or multicultural styles are considered to represent a combination of traditional and modern views. Examples of three major dimensions of traditional and modern styles are presented in Table 1.

The traditional-modern belief systems dichotomy was introduced by Robert Nisbet to conceptualize a conflict in societal ideology that emerged at the end of the 19th century. The conflict was given form by the democratic revolution and the Industrial Revolution. These worldviews define different perceptions of reality. Presently, the conflict has become more acute during the past two presidential elections and may be thought of as the genesis of dialectically opposed sets of values and goals. For example, the view of creation of the universe has been couched recently as intelligent design, or creationism, versus evolution. Another example is the perspective of the primary life goal, characterized by some as seeking salvation in the afterlife and by others as attaining meaning and fulfillment in the present life. Another way of understanding how cultural styles influence individuals is the measure of a "good" person. In some literature, it has been presented as emotional IQ and sense of community, and in other literature, it is presented as a more individual orientation to life.

Cultural styles are transmitted or moderated through socialization and the educational practices of parents, communities, and cultures. By the time children first encounter educational environments, they have developed preferred cultural styles that either match or do not match the preferred cultural styles of teachers, as well as the educational practices and programs of the schools they attend. The degree of match, in turn,

Table 1 Traditional and Modern Cultural Styles

	Type of Belief	
Focus of Belief	*Traditional*	*Modern*
Creation of the universe	Sacred	Rational
Primary life goal	Salvation	Living life to the fullest
Characteristics of the "good person"	Caring and compassionate "role model"	Self-determination or freedom to pursue life goals

determines the child's quality of adjustment and general success in learning environments. Flexibility in teachers and educational environments is essential to ensure student enthusiasm for learning and success in school.

Cognitive Styles

To ensure student success in educational settings, it is important to match their cultural styles and their personality styles. Educators are encouraged to become sensitive to the manner in which students approach new learning situations and to their motivation to learn. The type of relationship students prefer to have with teachers is also important, as is the types of rewards that encourage learning and retention.

Cognitive styles have five major components: relational styles, communication styles, motivational styles, learning styles, and teaching and supervisory styles. The three major cognitive styles identified by Ramirez and Castaneda are as follows:

1. *Field independent*—this personality style is characterized by a formal manner of relating to others (including both peers and teachers), a communication style that is impersonal and brief (economical use of words), a motivational style that is oriented toward achievement for the self, a learning style that is analytical and focused on the details of what is to be learned, and a teaching-supervisory style that encourages learning by discovery and trial and error.

2. *Field sensitive*—this style is characterized by a tendency toward a more personal and informal style of relating to others; a communication style that is more personalized and detailed; a motivational style that is driven by social rewards, especially praise from teachers and peers; a learning style that is more relational, with a tendency to merge knowledge from different fields or different sources; and a teaching-supervisory style that emphasizes modeling and mentoring.

3. *Bicognitive*—this last style is characterized by the ability to switch personality and cultural styles depending on the learning challenge or situation encountered. This style encompasses creativity as the ability to combine features and characteristics of both field independent and field sensitive personality styles to develop new approaches to problem solving. A bicognitive student has the ability to apply a combination of insights from different disciplines and the flexible use of analytical and relational thinking styles.

Some examples of some classroom behaviors related to field independent and field sensitive styles are presented in Table 2.

Research on development of individual differences in cognitive styles indicates that socialization practices play a key role in determining learning behavior preferences in children. The teaching styles of parents and other family members and the types of learning behaviors encouraged in families seem to contribute to the development of preferred learning behaviors in

Table 2   Learning Behaviors

Instructional Relationship to Teacher

Field independent	1. Likes to try new tasks without teacher's help
	2. Seeks nonsocial rewards such as gold stars
Field sensitive	1. Seeks guidance and demonstration from teacher
	2. Seeks rewards such as praise, which strengthen relationship with teacher

Thinking Style

Field independent	1. Focuses on details and parts of things
	2. Likes discovery or trial-and-error learning
Field sensitive	1. Focuses on the global nature of concepts of ideas
	2. Prefers discovery of concepts in humanized, personalized, and story format

children. For example, one family may emphasize the importance of people, such as how people relate to each other and their respective roles in the family. The child in this family may learn about motivation and about why people do certain things in certain ways. For this student, a close interaction between teacher and learner is encouraged. These children may learn many things by modeling what they see older people doing. These children develop the skills of what has been come to be known as emotional IQ.

Another family may encourage their child at an early age to find things independently. Children from such families may have learned to work out problems by trial and error. They will enjoy experiments with new materials and new games, but may not enjoy learning concepts requiring careful and exact observation and imitation. They may not learn to feel like they are part of a group as early as children from the family described previously.

Encouraging Development of Cultural and Cognitive Style Flexibility

The cultural and cognitive flex model applies the philosophy of cultural and individual democracy to educational environments. Horace Kallen introduced cultural democracy in response to the concern expressed by some segments of American society that some immigrants were slow at becoming Americanized. Observing a continuing persistence of languages, values, and lifestyles associated with nations and cultures of origin, Kallen suggested that it was necessary to adopt a cultural democracy perspective rather than a melting pot ideal for American society. That is, rather than insisting on forced enculturation to the American way of life, cultural democracy, derived from the idea of political democracy, would allow development of flexibility of beliefs and ways of life. It would give citizens the option to remain identified with their culture of origin while they adopt mainstream American culture. What Kallen proposed was an alternative to the melting pot; an opportunity for the development of bicultural or multicultural orientations to life.

Ramirez and Castaneda proposed that if cultural democracy were applied to education, it would facilitate the development of cultural and cognitive style flexibility in students regardless of ethnic origins. Specifically, cultural democracy was perceived as a philosophical precept that recognizes that the way a student communicates, relates to others, seeks support and recognition from his or her environment, and thinks and learns is a product of the value system of his or her home and community. When students are allowed to use their preferred cultural and cognitive styles, they achieve success and can proceed to learn their nonpreferred learning style. Through this process, students become bicultural and bicognitive. They become Americanized, but they also retain their identities with their original cultures. Individual democracy in the context of education, introduced by John Dewey, concluded that a society that constricted individuality would lose sight of democratic principles. He emphasized that public policy should respect whatever is unique and distinctive in each individual.

Research by Ramirez and Castaneda and also Barbara Cox and Ramirez found that most students have personality types that are bicognitive to differing degrees. Most learners, however, have a preference for one or the other style. By observing children in classrooms and college students in different learning environments and social and work settings, results showed that the degree of success in education and in life in general depends on their degree of bicognitive functioning. The researchers concluded that the ultimate goal of education was the promotion of flexibility in cultural and cognitive styles resulting in development of the total person.

To promote flexibility in cultural and cognitive styles, it is necessary to individualize instruction by encouraging flexibility in teachers and educational environments. Changes were suggested in four major components of instruction:

1. *Teacher training to enhance understanding of diverse cultural styles in American society.* Specifically, teacher training programs need to present the sociological, psychological, and anthropological characteristics of the different ethnic groups in this country. The information needs to be presented in an educational perspective. For example, cultures that are more traditional tend to teach by modeling and direct instruction, whereas more modern cultures generally teach by trial and error using a less directive approach. Teachers could be trained to use culture-matching teaching strategies and could learn how to assess the preferred cultural and cognitive styles of students. Teachers also need to perform self-assessment to become aware of their own preferred cultural styles and their relation to teaching styles.

2. *Development and selection of different kinds of curricula to facilitate learning among children with different cultural styles.* The principal focus of curriculum development should be on how new material is introduced and presented such that it incorporates cultural and cognitive styles. For example, if the lesson plan calls for teaching about rainbows, allowing children to tell stories about their experiences with rainbows is likely to be meaningful to those from traditional cultures. On the other hand, children from modern cultures might prefer experimenting with a bowl of water and a small mirror inside the bowl to capture sunlight to show how rainbows are formed. Both teaching strategies will benefit field sensitive/traditional as well as field independent/modern learners because they offer a diversity of teaching approaches that serves to match the preferred styles of each group while simultaneously providing the opportunity to add the nonpreferred skills to existing learning repertoires.

3. *Training teachers to use learning experiences and environments to teach respect for cultural differences and cross-ethnic cooperation.* This is particularly applicable in social studies curricula where the objective is to encourage students to understand familial, cultural, individual, and gender differences. For example, teaching about holidays celebrated by different cultures and discussing the lives of prominent women in society as well as the achievements of historical figures of different ethnic, religious, racial, gender, and socioeconomic backgrounds should be a part of social studies curricula. Teachers should also be taught to discourage use of ethnic and racial slurs and language that tends to disparage any group or person. Teachers should engage in proactive behavior to discourage ethnic isolation and should emphasize children of different genders and groups working together as a team to prevent isolationism and clique formation.

4. *Training for use of tests and other procedures for assessment of academic progress and intellectual ability to match the cultural styles of learners.* For example, children from traditional cultures are more likely to prefer essay-type questions or tasks that tap creative writing skills, whereas children from modern environments usually prefer multiple choice-type tests and assignment of projects that apply concepts to problems. The negative consequences of mismatch can have dire consequences for both learners and teachers because inaccurate conclusions can be drawn about the capabilities of students and the effectiveness of teachers.

Once instruction and curriculum content have been addressed, the students can then begin the match-mismatch process. The cultural and cognitive flex model of education follows three instructional steps: match of preferred cultural and cognitive styles, initiation of mismatch through match, and continued match and mismatch to promote flexibility and adaptability.

The first step acknowledges and respects the child's cultural and familial experiences in learning by matching the student's preferred cultural and cognitive styles. Assessment of the preferred styles of children when they first come to school is essential. Identification of the preferred cultural and cognitive styles of each child is done through classroom observation of behavior in an environment that allows for freedom of individual expression of cultural and cognitive styles. Once assessment is completed, the child is placed in one of three groups where instruction will be matched to preferred styles. This is the initial match.

The second stage is to gradually introduce mismatch of styles once the child achieves mastery in his or her preferred style. This is done by introduction to a bicognitive-bicultural group in which the nonpreferred styles are presented through use of the dominant styles; for example, introducing individual competition through group competition by asking group members to focus on which member achieved the most points during the exercise. This approach benefits the field sensitive and traditional learner who is in the process of becoming more bicognitive and bicultural.

Upon mastery in this environment, children are then introduced to the third stage in the flexibility development process, which is placement in a group where the child is introduced to new instructional strategies. These strategies include exposure to learning and testing materials that are written and presented in the previously unfamiliar styles. For example, a teacher introducing the nonpreferred style who uses the discovery approach might proceed in this manner: "Yesterday, I showed you how to find out if two triangles are equal. I have also showed you how to find out if two squares are equal. Now you know the shortcut I use in finding out the area of something. I have some rectangles for you to look at, and I want you to find out if they are the same, but I want you to do it in the way you think I would, using the shortcut I used with the triangles and the squares." After becoming familiar with the mixture of cognitive and cultural styles, the child is ready for the transfer to the group in which teaching and curriculum are based almost exclusively in the unfamiliar style.

The fourth and final stage is progression into an advanced mixed-styles group in which both styles are used to reinforce maximum flexibility and adaptability. An example of a lesson plan that would be presented in this group is as follows: In a field independent math lesson, a child whose preferred cognitive style is field independent and cultural style is modern feels familiar with an abstract, impersonal curriculum. This student enjoys individual competition and learns most advantageously when the teacher emphasizes individual effort. In a field sensitive and traditional curriculum (based on the child's unfamiliar styles), the child easily makes the transition from a math lesson that emphasizes inductive reasoning and abstract concepts to one that emphasizes deductive reasoning and personalized/humanized concepts. This student can work just as well cooperating with classmates as competing with them in the field independent-modern math lesson.

The flexibility and adaptability that students develop can transfer to areas of life outside educational environments. They are likely to become more receptive to other languages and other cultures, giving them the opportunity to achieve linguistic and cultural flexibility.

To succeed in preparing students for success and good psychological adjustment in a complex and technological society, we must get away from the one-size-fits-all mentality that is presently so much a part of the American educational system. A focus on individualizing instruction and utilizing technology can make this possible. For example, a computer program can determine a child's preferred cultural and cognitive styles by assessing the preferred styles at the start of the school year. The program would teach a new concept by asking simple questions regarding preferences for content and strategies for learning. The program could also monitor the progress of the student through the steps of cognitive and cultural flex achievement. Each individual student's record can become a part of his or her educational history and used to plan his or her advancement through each grade. In addition, it could serve as a useful record for different teachers when students change schools or move from elementary through middle and high school.

The work by Ramirez and Castaneda and that of their colleagues focused primarily on Mexican American children, showing how the traditional cultural styles of Latino children were being mismatched by the predominantly modern styles of the American public school system. Mismatch between preferred cultural and cognitive styles and the teaching styles, curriculum, and educational procedures of schools can result in mismatch shock, defined as exhibiting stress and feelings of failure and alienation. To this day, mismatch is the major reason why Mexican Americans continue to have the highest dropout rate of any ethnic group in American society. The mismatch factor is also important for the high failure rate in other ethnic/racial cultural groups in the United States who are being mismatched by educational institutions, teachers, and professors. As reported by the *New York Times* in February 2007, even when some schools make the effort to introduce diversity in curriculum offerings, school boards may stifle the effort. In fact, all school-aged children and college students, regardless of race, ethnicity, or socioeconomic class, are likely to be mismatched and to become discouraged and frustrated. A recent study by the Education Testing Service found that high school graduation rates and achievement gaps in reading and math skills have not changed very much in the past 20 years.

Assessment and Testing

As testing has become an indispensable component of the American educational system, it has also become a barrier to individualization of instruction and to a culturally democratic education. Cultural and cognitive styles in students can be mismatched by standardized administration procedures as well as mismatch of the content and structure of the assessment instruments. Mismatch through testing can lead to inaccurate conclusions regarding level of curriculum mastery and general intellectual ability. These two issues have become salient in the past few years because of the emphasis on testing related to the No Child Left Behind Program. Teachers have been forced to "teach to the test" as administrators focus on school test performance scores for teacher evaluations, salary increases, and recognition. This forces teachers to remain with a field independent type of curriculum and teaching strategies, thereby overlooking individual and cultural differences in their students.

Ultimately, there is a double threat of mismatch: first, between the teacher and student, and second, between the teacher and the test-oriented curriculum. When teachers are forced to "teach to the test," opportunities for collaboration between children and teachers on creative classroom projects are lost. In one extreme, creative science or art projects, which

encourage comparing classroom work to real problems, are considered secondary to rote and mechanized drills. For example, when teaching about the weather, students could be encouraged to play the role of meteorologists who predict and chart weather patterns rather than just memorizing information.

Regarding procedures of administration for tests and evaluation instruments, research has found that field sensitive and traditional children do best under conditions in which the administrator befriends them and presents the testing situation as a game rather than the more formal and distant conditions usually suggested for test administration. Additionally, children who are mismatched by testing procedures and instruments can be barred from promotion or graduation or incorrectly diagnosed as mentally retarded or learning disabled.

Would-be reformers of the American public education system propose a voucher system as a solution to the failure of the current system. However, the cultural and cognitive flex system of education would argue that the answer to educational reform does not necessarily lie with the privatization of education, but rather with the way teaching is usually executed, with what is emphasized in learning environments, and with the way schools are organized. The objective is to eradicate factory-model schools and the one-size-fits-all mentality. When students are provided with personalized, challenging, and meaningful experiences, they are allowed to use their preferred cultural and cognitive styles and gradually learn unfamiliar styles without negative consequences. Flex development takes place and conditions are established for lifelong enthusiasm for learning and social and intellectual growth.

Cognitive and Cultural Flexibility in a Diverse and Global Society

Adaptability and flexibility are necessary to the attainment of life satisfaction and successful participation in a complex, technological, diverse, and global society. Flexibility of learning is a necessity in our new society, as is adaptability to working with people from diverse backgrounds. In a leadership research study, leaders of ethnically mixed groups, who varied in cultural and cognitive flex, were asked to make every effort to reach group consensus on solutions to group problems that were impossible to solve (the group members were instructed by the experimenter to play obstructionist roles). These leaders were observed and rated on effectiveness of effort. The findings supported the hypothesis that leaders who were more culturally and cognitively flexible were significantly more likely to use effective behaviors such as establishing group cohesiveness, soliciting the opinions of all group members, and establishing a direction for the group than the leaders who had scored lower on flexibility measures.

Encouraging flexibility through education can also reduce dropout rates and promote multicultural participation, as revealed by the findings of a long-term research study that applied the cultural and cognitive flexibility model in a Southern California school. Students in the experimental and control classrooms were matched with respect to ethnicity and family socioeconomic status. The experimental participants were exposed to the flex program from kindergarten through third grade, whereas the control group students were taught under the standard curriculum of the same school. Former participants in this study were compared 36 years later according to various outcome measures. The findings indicated that the experimental classroom participants were significantly more likely to graduate from high school, to have multicultural friendships, and to participate in activities of the different cultural groups represented in their communities of residence. They also scored higher on the reading, spelling, and math subtests of a standardized achievement test.

The challenge to improve the American system of education facing parents, educators, and social scientists is enormous. As American society and its economy become more complex, the demand for well-educated, flexible, adaptable citizens becomes more pressing. The individuality of students and teachers needs to be acknowledged and recognized, and educational environments that recognize, respect, and promote this individuality need to be developed.

Manuel Ramirez III

See also Cognitive Development and School Readiness; Ethnicity and Race; Identity Development

Further Readings

Castaneda, A. (1984). Traditionalism, modernism, and ethnicity. In J. L. Martinez & R. H. Mendoza (Eds.),

Chicano psychology (2nd ed.). Orlando, FL: Academic Press.

Cox, B. G., & Ramirez, M. (1981). Cognitive styles: Implications for multiethnic education. In J. A. Banks (Ed.), *Education in the 80s: Multiethnic education.* Washington, DC: National Educational Association.

Ramirez, M., & Castaneda, A. (1974). *Cultural democracy, bicognitive development and education.* Orlando, FL: Academic Press.

COGNITIVE BEHAVIOR MODIFICATION

Cognitive behavior modification (CBM) is a blending of two conceptual models of management behavioral and cognitive theory, and as such is a powerful tool for changing trajectories of behavior in children and adults. One of the most widely researched and heavily evidenced types of therapy or intervention, CBM provides an evidence-based model for practice in public and private applications. Behavioral models of management originate with John Watson, Edward Thorndike, and B. F. Skinner and are based on learning theories of operant conditioning. Behavior is thought to be "learned" and as such can be "unlearned." All behavior serves a function for the individual, such as escape and avoidance or access to a reinforcer or reinforcement of some kind. Those reinforcers could be tangible or intangible things or conditions. Cognitive theory involves thoughts and feelings, two things that a behaviorist could not identify or measure overtly. Cognitive theorists would discuss cognitive structures and internal dialogue as the reason for behavior. This internal dialogue is also called *self-speech* or *self-talk* and is believed to be modifiable through self-instruction training, whereas a purely behavioral model would seek to change the antecedents and consequences that maintain the behavior.

Definition and Description

Cognitive behavior modification is the theory and practice that people's thinking about events, rather than the events themselves, are responsible for their actions, and that thinking can be modified and lead to behavior change. CBM involves overt behavior but also considers the verbal and internal processes that monitor and guide the more observable behavior. Interventions that are grounded in CBM include self-dialogue and thinking as component antecedents and consequences in changing behavior.

Behavioral Model: Observable and Measurable Events

Antecedent → Behavior → Consequence

Example: Jack is teased by Fisher → Jack hits Fisher → Fisher quits teasing

Jack has learned that hitting will result in escaping teasing and may be more likely to hit the next time there is an occurrence of something Jack wants to escape. A behavioral intervention would target teaching Jack to escape appropriately. This would reflect the function of the behavior (escape) while teaching a new skill to get the same need met.

Cognitive-Behavioral Model

Antecedent → Thinking/Belief System → Feeling → Behavior

Example: Jack is teased by Fisher → Jack thinks about getting teased and has a strong emotional response → The anger or frustration or embarrassment → Leads to hitting

The feelings and behavior in this model are the focus of the intervention rather than the behavior of hitting. A cognitive behavioral intervention would target the thinking and feeling or self-dialogue that occurred rather than attempt to control the consequences and antecedents.

Cognitive behavior modification is a form of intervention that emphasizes the important role of thinking in how people feel and what they do. Cognitive behavior modification involves the attribution of beliefs to people's thoughts that theoretically cause their feelings and behaviors. The benefit of this CBM model is that thinking and beliefs are conceptualized as learned. Thinking, feeling, believing (self-talk, self-narration, self-schema) as a learned behavior means people can change the way they think in order to feel or act, regardless of the situation. CBM can be thought of as a theory, a system of strategies, and a series of techniques. The theory is based on the idea that the processing of information is crucial for the survival of any person or individual.

Cognitive-behavioral therapists teach that when people's brains are healthy, it is their thinking that causes them to feel and act the way they do. Therefore, if a person is experiencing unwanted feelings and behaviors, it is important to identify the thinking that is causing the feelings or behaviors and learn how to replace this thinking with thoughts that lead to more desirable reactions.

Purpose

Theoretically, cognitive-behavioral therapy can be employed in any situation in which there is a pattern of unwanted behavior accompanied by distress and impairment. It is a recommended treatment option for a number of mental disorders, including affective (mood) disorders, personality disorders, social phobia, obsessive-compulsive disorder, eating disorders, substance abuse, anxiety or panic disorder, agoraphobia, posttraumatic stress disorder, and attention deficit/hyperactivity disorder (ADHD). Patients with sleep disorders may also find cognitive-behavioral therapy a useful treatment for insomnia.

Cognitive-behavioral therapy combines the individual goals of cognitive therapy and behavioral therapy. Pioneered by psychologists Aaron Beck and Albert Ellis in the 1960s, cognitive therapy assumes that maladaptive behaviors and disturbed mood or emotions are the result of inappropriate or irrational thinking patterns, called *automatic thoughts.* Instead of reacting to the reality of a situation, an individual reacts to his or her own distorted viewpoint of the situation. For example, a person may conclude that he or she is "worthless" simply because he or she failed an exam or didn't get a date. Cognitive therapists attempt to make their patients aware of these distorted thinking patterns, or cognitive distortions, and change them (a process termed *cognitive restructuring*).

Behavioral therapy, or behavior modification, trains individuals to replace undesirable behaviors with healthier behavioral patterns. Unlike psychodynamic therapies, it does not focus on uncovering or understanding the unconscious motivations that may be behind the maladaptive behavior. In other words, those who are strictly behavioral therapists don't try to find out why their patients behave the way they do, they just teach them to change the behavior.

Cognitive-behavioral therapy is used to mitigate maladaptive behaviors through the use of covert self-statements with the behavioral modification techniques of behavioral therapy. The therapist works with the patient to identify both the thoughts and the behaviors that are causing distress, and to change those thoughts in order to readjust the behavior. In some cases, the patient may have certain fundamental core beliefs, called *schemas*, that are flawed and require modification. For example, a patient suffering from depression may avoid social contact with others and suffer considerable emotional distress because of this isolation. When questioned, the patient reveals to his therapist that he is afraid of rejection, of what others may do or say to him. Upon further exploration with his therapist, they discover that his real fear is not rejection, but the belief that he is uninteresting and unlovable. His therapist then tests the reality of that assertion by having the patient name friends and family who love him and enjoy his company. By showing the patient that others value him, the therapist both exposes the irrationality of the patient's belief and provides him with a new model of thought to change his old behavior pattern. In this case, the person learns to think, "I am an interesting and lovable person; therefore, I should not have difficulty making new friends in social situations." If enough irrational cognitions are changed, this patient may experience considerable relief from his depression.

Interventions

Intervention in CBM works on the identification of one of two types of beliefs: rational and irrational. Several types of interventions work from this basic principle.

Cognitive restructuring is the process of replacing maladaptive thought patterns with more constructive thoughts and beliefs. Maladaptive thoughts are those considered unsuitable or counterproductive; for example, maladaptive behavior is behavior that is inappropriate to a given situation.

Verbal mediation is the process of identifying the self-talk or inner speech that will lead to goal fulfillment or task achievement. A script of self-talk that corresponds to the targeted or desired behavior is used in mental rehearsal to shape behavior.

Problem solving involves deconstructing problem scenarios and developing alternative causal explanations and inferences along with alternative solutions or behaviors. These alternatives are then mapped to

determine best choices or best options for subsequent conditions of the same problem.

Self-instruction is the process of teaching steps toward task completion. There are specific steps in self-instruction as follows:

1. *Demonstration by Model.* The teacher/therapist models behavior and says out loud what he or she is doing.

2. *Modeling With Overt Adult Guidance.* The student/client performs the task while talking to him- or herself out loud. The teacher/therapist corrects the student/client, helps with difficulties, and gives positive feedback.

3. *Modeling With Overt Self-Guidance.* The student/client performs the task while orally talking him- or herself through without guidance from the teacher/therapist.

4. *Modeling With Faded Self-Guidance.* The student/client whispers the self-guidance while performing the task.

5. *Self-Guidance.* The student/client speaks silently while performing the actions.

Tips for Implementation. Frequent practice is needed for fluency and mastery. Participants all need verbal rehearsal and opportunities for application. Start small before tackling large issues. Use visual prompts.

Benefits

CBM has been reported to have a host of benefits. Individuals report having more self-control, and some research reports its use for increasing time on task in classrooms, improved performance on intellectual tasks, decreases in impulsivity, and increases in attention. Other positive effects include an increased awareness of feelings and maladaptive belief systems, as well as an increase in self-evaluation of how behaviors related to personal behaviors.

CBM has demonstrated effectiveness for many children who benefit from modeling and rehearsing. CBM increases the attention of children after limited treatment exposures. CBM improves the processing abilities of aggressive students and increases problem-solving abilities. For students with ADHD, CBM has demonstrated effects for reducing the impulsiveness of many students by having them implement problem solving, and for students with learning disabilities,

anxiety disorders, depression, anger, and impulsivity, improved performance is expected from CBM.

Cognitive behavior modification is different from psychodynamic therapy, and they should not be confused one for the other. Psychodynamic therapy is a therapeutic approach that assumes that dysfunctional or unwanted behavior is caused by unconscious, internal conflicts and focuses on gaining insight into these motivations. CBM is a focus on the verbal behavior of an individual. It is covert but not unconscious. This verbal behavior is also conceptualized as automatic thoughts, which automatically come to mind when a particular situation occurs. Cognitive-behavioral therapy seeks to challenge automatic thoughts for their rational or irrational basis. Schemas are the fundamental core beliefs or assumptions that are part of the perceptual filter that people use to view the world. Cognitive-behavioral therapy seeks to change maladaptive schemas. Relaxation techniques are sometimes used to relieve stress. Exercise, biofeedback, hypnosis, and meditation are all effective relaxation tools. Relaxation techniques are used in cognitive-behavioral therapy to teach patients new ways of coping with stressful situations but are not a direct component of CBM.

Future research on CBM includes a need for data on the long-term effectiveness of using the procedures and the applicability for individuals with specific disabilities, such as attention deficits and emotional and behavioral disabilities across ages.

Denise Soares and Kimberly J. Vannest

See also Applied Behavior Analysis; Behavior Modification; Classical Conditioning; Cognitive View of Learning

Further Readings

Hersen, M. (Ed.). (2005). *Encyclopedia of behavior modification and cognitive behavior therapy.* Thousand Oaks, CA: Sage.

Kaplan, J. S. (1991). *Beyond behavior modification: A cognitive-behavioral approach to behavior management in the school.* Austin, TX: Pro-Ed.

Kazdin, A. E. (1982). Current developments and research issues in cognitive-behavioral interventions: A commentary. *School Psychology Review, 11*(1), 75–82.

Kendall, P. C. (1993). Cognitive-behavioral therapies with youth: Guiding theory, current status, and emerging developments. *Journal of Consulting and Clinical Psychology, 61*(2), 235–247.

Meichenbaum, D. (1977). *Cognitive-behavior modification: An integrative approach.* New York: Plenum.

Swaggart, B. L. (1998). Implementing a cognitive behavior management program. *Intervention in School and Clinic, 33*(4), 235–238.

COGNITIVE DEVELOPMENT AND SCHOOL READINESS

Many of those who remain committed to Jean Piaget's cognitive developmental theory see important implications for our conceptions of readiness for formal aspects of school learning. The general principle of strong correlations between measures of cognitive development and school achievement is complemented by Piaget's proposal of a hierarchical sequence of four cognitive developmental stages: sensori-motor, pre-operational, concrete operational, and formal operational stages. In many Western school systems, the timing of the movement of children from the informal learning settings of kindergarten and preschool to the first year of formal curriculum-based learning in primary or grade school coincides with the transition from pre-operational to concrete operational modes of thinking. Furthermore, the change in curriculum demands of theoretical mathematics, language, and science learning (for example) in high schools or secondary schools often occurs as students' minds are claimed to be in transition from concrete operational to formal operational thinking. However, whereas the timing of the school transitions is remarkably age-determined, the corresponding cognitive developmental transitions are merely age-related.

Three particularly important and widely implemented pedagogical outcomes of the study of cognitive developmental theory have been the play-based nature of preschool learning experiences, the hands-on approach to mathematics learning in the grade school years, and a problem-solving approach to high school science curricula. For almost two decades, two key U.S. journals, the *Journal for Research Into Science Teaching* and *Science Education*, were dominated by reports on the relationships between high school science learning and formal operational thinking. Researchers such as Anton Lawson, Warren Wollman, and Robert Karplus were foremost among researchers of that era. Although many would observe that the Piagetian influence in science education is well past its zenith in the United States, that is not the case in the United Kingdom, and that is certainly not the case for preschool education. The remarkable success of the team of Michael Shayer in the United Kingdom in applications to high school science pedagogy, in particular, and education in general, shows the benefit of theoretically informed long-term commitment. Preservice education of early years teachers remains under the sway of the Piagetian oeuvre and the early influence of Constance Kamii has been continued by others, such as Rheta Devries and Betty Zan.

Readiness for High School Science

Shayer and part of the Concepts in Secondary Mathematics and Science team at the University of London devised a cognitive demand taxonomy directly from the Barbel Inhelder and Jean Piaget book on the development of formal thought and used it to estimate the intellectual demands of the U.K. national high school science curriculum. Their subsequent representative national survey of children's cognitive development revealed a pedagogically important mismatch between the cognitive demands of the science syllabus content and the cognitive developmental capacity of high school students in the United Kingdom to understand that material. Therein lies the conceptualization of the relationship between cognitive development and lesson content as it relates to school readiness: The cognitive demands of the formal curriculum exceed the cognitive capacity of many students who have been promoted into that new school setting on the basis of some age-determined criterion. Two alternate resolutions for this mismatch seem possible in order to address the readiness for learning issue: reduce, reorganize, or reschedule the cognitive demands of the curriculum for all/many/some learners; or promote the students on the basis of cognitive readiness for learning rather than by age. In the face of intransigence of curriculum experts, or, in the U.K. situation, the demands of university entrance exams, a third, much more radical, solution seemed appropriate to Shayer's team. They developed the *Thinking Science* intervention to maximize the development of formal operational thinking structures in middle school and junior high school students. *Really Raising Standards* reported that not only were cognitive development rates improved, but students who received the

special thought-provoking lessons later outperformed controls in science, mathematics, and language on the national public examinations used to decide access to universities. Lorna Endler found similar results for her replication of the project in Australia, which prompted a district-wide Goals 2000 implementation in Oregon in the United States. Stanbridge's research on her own radical constructivist science teaching over almost a decade revealed that understanding of core physical science concepts was constrained by the level of cognitive development of the learner; a cognitive ceiling on achievement was implied by Panizzon's research with high school, college, and university students.

Readiness for Year One

Birthday-related entrance into formal, year one learning is almost ubiquitous in Western schooling systems. Cognitive developmentally informed investigations into school performance in the early years yield three results that are quite consistent with the findings from the high school research above. The summaries provided in Bond derive from a sequence of studies in which research students measured both students' levels of cognitive development and their school achievement using either mandated or teacher-made assessments. First, cognitive development is remarkably heterogeneous for any school group selected by age or year of schooling. Second, curriculum requirements are out of the grasp of many students based on their Piagetian level of cognitive development. Finally, academic success at grade/primary school clearly is related to that development. The general findings are that academic success is apparently more dependent on cognitive development than on the age (or year level) of the child. This implies a consequently smaller role for curriculum and learning experiences in school achievement. The wide range of achievement for any age cohort is consistent with the huge variability in stage of cognitive development in any age-determined class group.

These generally consistent results prompted Maley to address directly the question of the preschool to year one transition and the roles played by cognitive development and the social/psychomotor indicators often recorded by preschool teachers. Predictably, children at the end of their preschool year found it much more difficult to provide conservation of number judgments with appropriate operational justifications

than they did to fulfill the requirements of their teachers' readiness checklist. Moreover, year one teachers were measurably less likely to judge children in the newly arrived year one cohort as "nearly ready" or "ready" for formal learning than were their preschool teachers just before the summer break. Tellingly, from the point of this encyclopedia entry, the indicators of concrete operational thinking were closely aligned to the readiness judgments of the year one teachers, suggesting a key role for considering cognitive development when assessing readiness, along with a need for preschool and year one teachers to construct a shared understanding about the "readiness" construct.

At the level of preschool education, Rheta DeVries and Betty Zan also propose important consequences for children's moral development that derive from a constructivist account of education derived from the theory and research of Piaget. They argue against more traditional approaches for the moral education of young children.

Trevor G. Bond

See also Measurement of Cognitive Development; Piaget's Theory of Cognitive Development

Further Readings

Adey, P., & Shayer, M. (1994). *Really raising standards: Cognitive intervention and academic achievement.* London: Routledge.

Bond, T. G. (2001). Ready for school? Ready for learning? An empirical contribution to a perennial debate. *Australian Educational and Developmental Psychologist*, *18*(1), 77–80.

Bond, T. G. (2003). Relationships between cognitive development and school achievement: A Rasch measurement approach. In R. F. Waugh (Ed.), *On the forefront of educational psychology* (pp. 37–46). New York: Nova Science.

DeVries, R. (1984). Developmental stages in Piagetian theory and educational practice. *Teacher Education Quarterly*, *11*(4), 78–94.

DeVries, R., Zan, B., & Hildebrandt, C. (2002). Issues in constructivist early moral education. *Early Education and Development*, *13*(3), 313–344.

Kamii, C. (1982). *Number in preschool and kindergarten: Educational implications of Piaget's theory.* Washington, DC: National Association for the Education of Young Children.

Lawson, A. E., Karplus, R., & Adi, H. (1978). The acquisition of propositional logic and formal operational

schemata during the secondary school years. *Journal of Research in Science Teaching*, *15*(6), 465–478.

Shayer, M., & Adey, P. (1981). *Towards a science of science teaching.* London: Heinemann.

Wollman, W. T., & Lawson, A. E. (1978). The influence of instruction on proportional reasoning in seventh graders. *Journal of Research in Science Teaching*, *15*(3), 227–232.

COGNITIVE VIEW OF LEARNING

Cognition refers to thinking and the mental processes humans use to solve problems, make decisions, understand new information or experiences, and learn new things. *Cognitive views of learning* focus on cognitive variables affecting learning—what goes on in people's minds before, during, and after learning. For example, when students are preparing to read a textbook chapter, they might think about what types of learning strategies they want to use to learn the important material in the chapter. They might decide to take notes and organize the material as they are reading, or to paraphrase each section as they read to see if they understood the material. Thinking about what learning strategies to use before starting to read may help students to read and learn more effectively. During the time they are reading, students can use the learning strategies they picked and check to make sure their strategies are working and that they are learning the material. After reading the text, students could think about what they did and decide if this was a good way to learn the material in this textbook. If it was, they might decide to use these methods again. If not, they may want to try something new when they read another chapter in this text. This example captures the essence of the cognitive view of learning—the most important part of learning goes on in the learner's mind. Studying these cognitive processes and fostering their development in children and adults is a major focus of educational psychology.

Cognitive views of learning have expanded both the breadth and depth of how people understand learning and the interactions among learners, teachers, and instructors; the learning environment; and learning materials. Not all students learn the same way, use the same methods, or interpret the environment identically. These individual differences or preferences for ways of thinking and learning have led educators to modify their teaching methods so that they are better able to provide support for individual students. Research on individual differences has led to the development of models of learning that focus on different types of cognitive processes and how they interact to produce meaningful learning rather than simple memorization. The focus of meaningful learning is on understanding and creating relatively long-term clusters of related knowledge and skills.

Models of how students process new information and experiences have evolved over time into complex, interactive cognitive models of learning. These models focus on goal-directed learning that involves the intentional use of cognitive, metacognitive, motivational, emotional, and self-management strategies and methods to reach learning and achievement goals. Successful learners think about and set goals for what they want to achieve, which helps them target and direct their use of learning strategies and methods. Setting and thinking about goals can also enhance motivation and commitment to learning. Using cognitive learning strategies involves building bridges between what the student already knows or has experienced and what he or she is trying to learn. For example, comparing and contrasting the political and economic causes of the Vietnam war with the Iraqi war can help students understand the Iraqi war. Thinking about what they experienced when they flew on an airplane can help students understand the physics of flight.

Metacognition refers to thinking about thinking. It is a metaprocess, that is, a process that goes along with another process to support it in some way. For example, solving a math problem involves cognition. Thinking about which cognitive processes to use to solve the math problem involves metacognition. Because educational psychologists know how critical it is to use effective cognitive strategies for learning, the importance of metacognition for learning is widely recognized. Motivation and emotion also have a significant impact on learning and have strong cognitive components. Thinking about how an individual can use something he or she is learning now or in the future can increase his or her motivation for learning. Keeping a positive attitude toward learning can also help students stay on task and not procrastinate.

The last area, self-management, involves planning, monitoring, regulating, and evaluating how one goes about learning. Students can benefit from using a systematic approach to learning. Basically, this involves setting a goal; selecting the cognitive learning strategies and other methods to use to learn the material;

implementing the methods selected; monitoring and evaluating the learning process to make sure progress is being made toward the goal; modifying or replacing strategies if they are not contributing to goal progress; and, at the end, evaluating the entire process to see if it is reasonable to try to complete a learning task like this the same way in the future or if it would make more sense to try different strategies and methods. Using a strategic approach to learning helps students build up a set of routines—sequences of thoughts and actions—that is effective for them so that they do not have to "reinvent the wheel" every time they try to learn something. For example, many people have established routines for how they read a favorite magazine and do not have to think about their goal or what learning strategies to use each time they read a new issue. In the same way, students who are being strategic should not have to think about how to read a chapter of text every time they read the same textbook, or when they are taking notes in class or completing a multiple-choice test.

Building on cognitive views of learning, educational interventions have been developed to help students succeed at all levels of formal education, industrial education and training, and everyday learning. This is a key area of research and work for many educational psychologists.

Claire Ellen Weinstein and Taylor W. Acee

See also Learning; Learning Strategies; Memory; Metacognition and Learning; Social Learning Theory

Further Readings

Boekaerts, M., Pintrich, P. R., & Zeidner, M. (Eds.). (2000). *Handbook of self-regulation.* San Diego, CA: Academic Press.

Zimmerman, B. J., & Schunk, D. H. (Eds.). (2001). *Self-regulated learning and academic achievement: Theoretical perspectives* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.

COMMUNICATION DISORDERS

The ability to communicate using speech is one of the most effortless, taken-for-granted human faculties. Its importance, however, becomes readily apparent when an individual fails to acquire the skill, loses the skill, or cannot use it effectively. Hearing and speech are inextricably connected: Development of language and speech depends on the exposure to spoken language during the early years of life. The range of speech communication disorders that arise during a lifespan are many, and they fall in the province of the field of speech-language pathology and audiology. Speech communication is the major focus of these fields. However, speech-language pathologists also give due importance to other modalities of expression—writing and signing—in their work.

Speech communication can be viewed as transmission of a message from a speaker to a listener. The linguistic encoding and transmission of a message by a speaker is a complex process; the reception and recovery of the message by a listener is an equally complex process. The field of communication sciences (speech and hearing sciences) provides empirical and theoretical underpinnings for the field of communication disorders. These fields are truly interdisciplinary in nature, drawing on wide-ranging fields in the physical, biological, and psychosocial sciences. Speech-language pathologists (SLPs) and audiologists diagnose and treat (manage) a variety of speech communication disorders in multiple settings, such as hospitals, schools, and nursing homes. To get a brief introduction to a sample of communication disorders that a certified SLP and audiologist is called upon to diagnose and treat (manage), the communication process must be considered in some detail.

Speech Production Process

Generally, three stages in the central brain processes are posited prior to production of speech. Each stage can be conceived of as a neural activation pattern that evolves and morphs at a fast pace into a pattern associated with the succeeding stage. The foregoing description suggests that the stages are organized serially, that is, processing in one stage is completed before the processing in the succeeding stage begins. However, there is evidence to suggest that processes in the stages and within each stage are part serial and part parallel. Thus, for example, fitting words in proper grammatical slots in a sentence does not occur one word at a time from beginning to the end but in parallel. This view is supported by a variety of slips of tongue data (e.g., **Intended**: *The map is in the car.* **Produced**: *The car is in the map*).

The first stage is thought processes leading to message formulation. Thought is language-neutral and, by symbolically representing thought using the vehicle of language, the message is made available to the conscious mind. This prelinguistic stage is generally conceived to be beyond the purview of communication sciences.

In the second stage, the message undergoes linguistic formulation. Linguistic formulation involves multiple processes such as selection of language to be used (if a speaker knows more than one language and the listener is equally competent in that language), the retrieval of words from the lexicon—long-term storage—to be fitted in a preselected sentence structure (syntactic frame) that specifies the linear order of parts of speech (grammatical categories) and phonological structure (i.e., how the sentence will sound).

The third stage is conceived to be motor planning of the linguistically formulated message. Motor planning is specification of timing and extent of muscle group contraction-relaxation in the peripheral organs involved in speech production. These organs, collectively referred to as the *speech production system*, compose the respiratory system, the laryngeal system, and the oro-nasal-pharyngeal system. The patterned muscular contraction-relaxations result in movement of ribs, lungs, vocal folds, velum (soft palate), jaw, tongue, and lips, all geared toward controlled expulsion of air and generation of sound using vocal folds, which in turn are shaped into a nearly concatenated series of overlapping sounds called speech. The motor plan reflects the sound- and word-specific movement requirements. Additionally, the motor plan also encodes for the effects the speaker intends the message to have on the listener. Thus, a speaker may state, persuade, seek to clarify, or seek clarification. These goals change the movement underlying a word sequence and hence the sound output. These changes are called *prosodic changes*. Prosodic changes can be illustrated by an example. A speaker may state "This book is interesting" or stress the word "This" to emphasize that this book, and not some other book, is interesting. In addition, the speaker may want to communicate enthusiasm, disgust, anger, irony, or approval. These affective components also contribute to changes in global and local acoustic features of the produced sentence. These changes are collectively referred to as *paralinguistic changes*.

A message repeated twice might not involve exactly the same motor plan. Motor plans are altered to respond to the current constraints on the speech mechanism. To take a simple example, to produce /p/ in the word "pat," the instructions to approximate the lips would differ depending on the current position of the lips. If the mouth is currently open, the muscles of the jaw as well as the lips will be involved in lip approximation. On the other hand, if the mouth is closed, the same target sound is achieved without the activity of jaw muscles. Peripheral sensory or afferent inputs inform central motor planning mechanisms about the existing constraints. In normal speech production, a rapid contextual adjustment is the rule rather than the exception. The use of different muscle groups to achieve the same acoustic/perceptual target is called *motor equivalence*. Motor equivalence underscores the principle that the brain uses variable means to achieve invariant ends.

The most important sensory signals arise from tactile (sense of touch) and proprioceptive mechanisms in the speech production system. They inform the motor planner about the constraints prior to initiation of speech. These feedback signals are supplemented with auditory feedback signals once speech begins.

Peripheral feedback is crucial during the early years of speech and language acquisition; their role in adult skilled speakers is still a matter that needs to be settled. Do skilled speakers depend on them all the time to monitor the progress of an utterance? There is ample evidence to show that they play a role in getting the production back on track if there is a sudden, unexpected perturbation to the speech production system. The relative importance of the type of feedback signal—auditory, tactile, or kinesthetic—is also an open question.

Respiratory, laryngeal, and oro-naso-pharyngeal structures have dual life-supporting functions: breathing and eating. Quiet breathing is automatic and governed by the respiratory centers in the medulla oblongata. During quiet breathing, durations of inspiration and expiration are nearly equal. While speaking, however, the duration of expiration is considerably increased in relation to the duration of inspiration. This change represents an active modification of expiratory muscle activity to generate a steady expiratory stream as a power source for speech.

The vocal folds in the larynx (commonly referred to as voice box) move to the midline to impede the expiratory breath stream. When the air pressure below

the folds exceeds air pressure above the folds in the vocal tract (pharynx, oral and nasal cavities), the folds are forcefully opened and set in a periodic to-and-fro movement, generating a voice. By changing the degree of vocal fold approximation (or degree of impediment to the flow of air) and the folds' length and tension, a speaker can generate degrees of emphasis (stress) and pitch to achieve linguistic, prosodic, and paralinguistic goals.

The voice (sound) is transmitted and modified by changing configuration of the vocal tract. The various cavities of the vocal tract act as acoustic resonators. Continuous speech involves rapid changes in vocal tract configuration, which in turn engenders rapid changes in resonating characteristics. The changes in resonating characteristics, in large measure, underlie the distinction between serially organized speech sounds produced during continuous speech. Changes in vocal tract configuration are achieved by the movement of lips, jaw, tongue, and velum.

Language Processing and the Brain

It has been known that the left hemisphere of the brain is dominant for language processing in most individuals ever since Paul Broca's description of a patient with a lesion in the left inferior frontal cortex region of the brain. The patient had effortful speech with relative preservation of comprehension skills. The clinical data from patients with lesions in the left hemisphere show that there is a degree of modularity—the frontal lesions generally affecting flow of speech while preserving comprehension, whereas posterior lesions in the parietal lobe affect comprehension skills while preserving flow of speech. There have been considerable attempts to relate symptoms with sites of lesions, with some degree of success. However, correlating symptoms with sites does not reveal how a brain functions in a healthy individual when engaged in various language tasks. The glimpses of complexity of neural activity when the brain is engaged in such simple language processing tasks as naming pictures and matching an appropriate verb with a stimulus noun are being highlighted by modern neuroimaging techniques such as functional magnetic resonance imaging, positron emission tomography, and magneto-encephalography. Such techniques reveal that there are multiple regions of activations within the left language hemisphere and the nondominant right hemisphere.

Disorders arising during linguistic formulation are collectively termed *aphasia* or *dysphasia.* Thus, a patient may have a problem in retrieving lexical items or words (anomia) from the lexicon, a problem with syntactic processing (syntactic aphasia), or a problem in sequencing the phonemes (apraxia) that constitute individual words. Aphasia typically results from cerebrovascular accidents such as hemorrhage, embolism, and ischemia in the left hemisphere. Children may exhibit language formulation issues even without brain lesions.

Hearing and Speech Perception

The process of recovery of encoded messages starts in the peripheral hearing mechanism (ear). The ear has three distinct parts: external, middle, and inner. The acoustic energy in the ear canal is converted into mechanical energy of the intricate motion of the ear drum and the three middle ear ossicles (bones). This motion in turn produces vibration of fluid in the inner ear. The motion of the inner ear fluid results in the motion of the basilar membrane. The end organs (hair cells) of the auditory nerve are located on the basilar membrane. Different regions of the basilar membrane are responsive (resonate) to different frequencies. Hence, the inner ear activity results in spectral (frequency by intensity) analysis of speech in terms of patterned excitation in the auditory nerve. The patterned excitation reaches the primary auditory cortex in the temporal lobe of the brain where sound is sensed. The secondary auditory cortex surrounding the primary auditory cortex is involved in putting the analyzed signal back (fusion) to recover the literal message and then the intended message.

The aforementioned thumbnail sketch of the hearing process provides a strictly serial view of hearing and speech perception. However, speech perception, defined as recovery of the sequence of phonemes constituting an utterance, is anything but serial. At any given point in time, a listener is not only analyzing the speech but also bringing to bear on the analysis contextual and knowledge constraints to arrive at a decision regarding sound and word identity. Thus, speech perception involves both bottom-up processes (acoustic input analysis) and top-down processes (contextual and knowledge constraint inputs to the analysis). There is considerable evidence supporting the idea that as a listener begins processing a spoken

word, the initial input (beginning of the word) activates a cohort of candidate words with that beginning. But as more and more of the spoken word unfolds, all the words are rejected (deactivated) in favor of the one the speaker intended. This is the idea behind Internet search engines with a search history—as an individual types a few beginning letters of a word, the system presents multiple candidate words beginning with the letter sequence. As the individual types more letters, more and more candidate words are rejected in favor of one the individual had used previously.

Communication Disorders in Educational Settings

Hearing Loss, Speech, and Language Development

Communication disorders can arise at any level of the communication process outlined above. An additional complication arises when communication disorders in educational settings are considered. Preschool and elementary school years are the years when children grow at a rapid pace physically, cognitively, and linguistically. There is a strong reason to believe that human children are biologically equipped to extract language from speech to which they are exposed. In essence, growing children implicitly acquire a system of rules (grammar) that will enable them, in Noam Chomsky's words, to understand and produce sentences that they have never heard before or produced before. However, there seems to be a critical period—roughly between birth and preadolescent years, and particularly between birth and 5 years—in which a child needs to be exposed to a language. A child's growth in other departments—cognitive, affective, and neuromuscular—must go hand in hand for a child to be an age-appropriate communicator.

The acquisition of language may be delayed in an otherwise normal child because of frequent middle-ear infections that result in temporary or permanent loss of hearing. The effect of frequent infections depends on whether the condition is unilateral or bilateral. Unilateral conditions typically do not affect language development; bilateral conditions do. The degree of hearing loss will have a direct impact on the phonological development—mastery of language-specific sound distinctions and word-specific changes that signify tense, number, and so on. A serious deficit in the middle-ear transmission function can result

in a 30- to 50-decibel (dB) sound pressure level (SPL) loss in the speech signal strength. Normal face-to-face conversation is typically around 60dB SPL. In this condition, the child is barely hearing speech and is missing most of the consonants, which are shorter and less intense than vowels. These phonological problems can be remediated effectively by addressing the hearing problem—using a hearing aid to amplify sounds—and with speech therapy that creates awareness of sound distinctions and direct speech therapy.

The extreme case is children who do not hear at all due to diseases—congenital or acquired—that affect the inner-ear processing. There are degrees of sensorineural hearing losses, and accordingly, one can expect degrees of language delays and impairment. A severe inner-ear impairment is generally recognized by parents or a caregiver; less severe conditions may receive delayed identification. In these conditions, early diagnosis has a direct impact on the effectiveness of rehabilitative procedures. Teachers and caregivers must seek professional help if they suspect hearing loss. An ear, nose, and throat specialist or an audiologist is the most appropriate person to contact. Early identification is so critical for the development of speech and language that there is a need for infant hearing screening programs.

Specific Language Impairment

Specific language impairment (SLI) is a condition in which children are normal in all respects and have adequate exposure and opportunity for learning a language but still fall significantly behind their peers. The prevalence of SLI in the general population is around 7%. An SLI child scores normal or above normal on nonverbal intelligence tests and does not have hearing loss, brain damage, or any other physical abnormality that can explain significantly low performance on language tests. A few recover, but many continue to suffer. SLI is significantly correlated with reading impairment.

Most English-speaking SLI individuals exhibit a deficit in using bound morphemes that indicate past tense (walk-walk*ed*), more than one (book-book*s*), and done by someone (climb-climb*s*). Some investigators have suggested that an SLI child has difficulty processing brief, rapidly unfolding sounds as an explanation for language impairment. This explanation suggests that, at its roots, it is a perceptual problem.

Several lines of evidence suggest a genetic basis for the disorder: More males than females have SLI, there is aggregation of SLI in families at a rate beyond the general population rate, and there is greater concordance (i.e., both suffer) among identical twins rather than fraternal twins.

There has been an in-depth investigation of a family (KE family) that had significant aggregation of individuals with SLI. The aggregation pattern in a three-generation pedigree indicates a strong autosomal dominance pattern. The genes responsible for this particular family's disorders have been mapped to a particular area on chromosome 7, and genetic analysis of an unrelated patient who exhibited similar language problems revealed a mutation in this same region. Although further studies are necessary, it may be that this mutation leads to a lack of a particular protein that affects the development of neural structures important for speech and language, the way similar proteins have been shown to influence neuronal development in other organisms. Subsequent study using neuroimaging techniques confirmed anatomical abnormalities in certain localized regions of the brain. The imaging study involved comparisons of brain anatomy of impaired and unimpaired members of the family. The overall genetic pattern exhibited by the KE family, especially with the discovery of the gene deletion, strongly suggests that a gene, or a small set of genes, has a major impact on language development.

Current studies are attempting to classify SLI patients into distinct subgroups based on the details of their impairment so that testing can determine more precisely the causes of the impairment.

Fluency Disorders

There are a host of disorders that affect the smooth flow of speech during early childhood. The most notable among them is developmental stuttering (DS). DS is a perplexing problem generally considered to be triggered in postlinguistic formulation stages.

Developmental stuttering, as the name suggests, manifests itself during the early years of rapid physical, cognitive, and language development. It is characterized by involuntary intermittent disruptions that take the form of fragmentation and repetitions of the beginning of words, prolongation of the initial sound of the word, or inappropriate silence or blocks. The onset is typically gradual and usually occurs when children start putting words together in sentences. The lifetime incidence of the disorder is estimated to be 5%, with a 1% prevalence rate in the population. Because the incidence of stuttering decreases precipitously from 2 to 10 years, the disparity in lifetime incidence and prevalence suggests that many children (around 80%) recover spontaneously, and a few continue to stutter into adolescence and adulthood.

The current understanding of the disorder is that it is etiologically complex, that is, both nature (genetics) and nurture (environment) play a role. The arguments in favor of a genetic basis are that it is more prevalent in boys than girls, the prevalence of stuttering in families of stutterers is significantly more than in the general population, and the concordance for stuttering among identical twins is very high compared to fraternal twins. Recent studies have shown that an autosomal dominance model, influenced by the gender of the affected individual and the affection status of the parents, is the best fit for the aggregation pattern in families. Recent neuroimaging studies have shown significant anatomical and functional differences between the brains of adult stutterers and normal speakers. The studies are still preliminary.

For a long time, stuttering was considered to be caused by anxiety, or it was considered to be something a person could overcome if only he or she tried. As a disorder, it was imitated, with negative consequences for individuals who were being imitated. The current understanding is that at its inception, it is purely a speech motor problem, and that anxiety as a factor starts having an impact on the basic problem once the disorder develops. Many researchers are focused on finding how children who spontaneously recover differ from those who do not.

Stuttering in school-age children is a big concern. The effect of this disorder on self-concept, coupled with generally negative reactions from others in the child's immediate environment, alters the child's psychosocial life. Later, it may compel him or her to limit his or her choices in terms of vocation and avocation. It is essential that teachers seek an SLP's help, and the child should be helped in the context of a team involving an SLP, teacher, and parents.

Dysarthrias

There is a large group of disorders that results from neuromuscular impairment that may affect the speed,

range, direction, strength, and timing of movements. These are due to either congenital or acquired lesions of the central nervous system. The speech symptoms can be mild or severe depending on the site and extent of the lesion. Children with dysarthrias exhibit imprecise articulation (sound production), an inability to control voice parameters such as loudness and pitch, and mis-timed speech movements. Dysarthrias occur independent of language processing disorders. Because the degree of impairment varies considerably, a thorough speech mechanism evaluation is a must to identify the problem and define the subtype of dysarthria before intervention can be planned.

Professional Training

This entry has highlighted a few communication disorders from a vast list. The field is enormous, and each disorder has an area of specialization. Diagnosis and treatment of communication disorders requires extensive training (a 4-year undergraduate program with communication sciences and disorders as a major) and a 2-year master's program followed by a year of clinical externship. Speech and language pathologists and audiologists must be certified by the American Speech-Language and Hearing Association before beginning to practice.

Nagalapura S. Viswanath

See also Language Disorders; Speech Disorders

Further Readings

Chomsky, N. (1964). *Current issues in linguistic theory.* The Hague: Mouton.

Fisher, S. E., Vargha-Khadem, F., Watkins, K. E., Monaco, A. P., & Pembrey, M. E. (1998). Localisation of a gene implicated in severe speech and language disorder. *Nature Genetics*, *18*, 168–170.

Foundas, A. L., Bollich, A. M., Corey, D. M., Hurley, M., & Heilman, K. M. (2001). Anomalous anatomy of speech-language areas in adults with persistent developmental stuttering. *Neurology*, *57*(2), 207–215.

Gleason, J. B. (1997). *The development of language* (4th ed.). Boston: Allyn & Bacon.

Laver, J. D. M. (1980). Monitoring systems in neurolinguistic control of speech production. In V. A. Fromkin (Ed.), *Errors in linguistic performance: Slips of the tongue, ear, pen and hand.* New York: Academic Press.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation.* Cambridge: MIT Press.

Owens, R. E., Metz, D. E., & Haas, A. (2007). *Introduction to communication disorders—A lifespan perspective* (3rd ed.). Boston: Pearson.

Raphael, L. J., Borden, G. J., & Harris, K. S. (2007). *Speech science primer—Physiology, acoustics, and perception of speech* (5th ed.). Phililelphia: Lippincott Williams & Wilkins.

Viswanath, N., Lee, H. S., & Chakraborty, R. (2004). Evidence for a major gene influence on persistent developmental stuttering. *Human Biology*, *76*, 401–412.

Watkins, K. E., Vargha-Khadem, F., Ashburner, J., Passingham, R. E., Connelly, A., Friston, K. J., Frackowiak, R. S., Mishkin, M., & Gadian, D. G. (2002). MRI analysis of an inherited speech and language disorder: Structural brain abnormalities. *Brain*, *125*, 465–478.

COMPETITION

Competition constitutes a primary facet of life, as individuals within any species strive to satisfy their needs for survival, reproduction, and successful rearing of offspring. For humans, social competition between genetically unrelated individuals of similar age is particularly salient in cultures with formal schooling and public institutions where individuals cooperate apart from families. Drawing on the child development and animal behavior literatures, this entry discusses the influences of biology, sex, culture, social structure, and age on competition between unrelated peers.

Biological substrates of competitive behavior in humans that have been identified include testosterone, cortisol, and alpha-amlyase. For example, in human males, testosterone increases in response to a competitive challenge. Winners' testosterone levels increase, whereas losers' levels decrease. Higher testosterone levels then promote positive expectations of future success.

Biological sex as well as gender identity, whether one feels male or female, relate systematically to competitive behavior. Most peer competition occurs between individuals of the same sex who compete for the same goals. Cross-culturally, beginning in the first years of life, females utilize more polite speech and nonverbal communication, such as smiling and behaving modestly, which serve to diminish retaliatory responses from potential competitors. In contrast, males more frequently exhibit behavior that invites retaliation, such as physically provocative behavior and verbal attempts to

dominate others through boasts, commands, threats, name-calling, and derisive jokes.

By middle childhood, males display more interference competition in which one individual must overtly prevent another individual from obtaining a desired outcome. In contrast, females find interference competition aversive and prefer to compete through more indirect or subtle means. Improving personal performance without reference to others; denigrating an absent competitor's reputation; or employing discreet nonverbal signals of competition, such as negative facial expressions, constitute female forms of competing. By middle childhood and increasingly in adolescence and adulthood, in their interests and leisure activities, males are more likely than females to engage in overt interference competition, frequently in the form of playing or observing competitive sports, often between unfamiliar players. In contrast, females derive more pleasure and relaxation from engaging in conversations with friends and relatives about comparisons of familiar individuals' levels of performance.

Results fit squarely within an evolutionary framework in which female mammals, who bear greater responsibility than males for their offsprings' survival, must avoid physical harm and maintain loyal social contacts in order to enhance their reproductive success. Those females who engage in competitive behavior risk retaliation which could endanger nearby offspring as well as their own ability to produce future offspring. By contrast, males, who vary more than females in reproductive success, benefit more from direct competition for mating opportunities. A male who continually attempts to outperform his male peers will hold a reproductive advantage and sire more offspring.

Local cultural norms also exert a powerful influence on absolute levels of competitiveness. Although few differences exist in early childhood, by middle childhood, children from differing countries and from varying cultural backgrounds within the same country differ in degree and form of competitive behavior. These same differences apply to cooperative behavior, and willingness to compete and cooperate frequently correlate positively, not negatively. Although scarcity of resources increases competition, when resources are plentiful but easily monopolized, this also enhances competitiveness.

Social structure also affects competitiveness. Mammalian species differ in the forms of social structures they create, such as large groups, smaller cliques, pairs, or solitary formations. Groups clearly enhance competitive behavior not only because the proportion of resources per individual declines, but also because the intensity of conflicts between individuals within a group can be regulated by third-party mediators, allies, the presence of alternate partners, and loyalty to the larger group. By contrast, conflicts between two individuals are marked by higher intensity, greater probability of harm, and enhanced likelihood of disintegration of the relationship.

Generally, human males more than females interact in larger groups, whereas females prefer one-on-one relationships. For both sexes, higher levels of competition occur in groups and lower levels between individuals. Males become engaged in intergroup competition, which intensifies as the number of individuals in each group increases. Groups likely accentuate males' competitiveness, whereas females' preference for one-on-one relationships diminish it.

Development also influences competitive behavior as children become more explicitly conscious of their own interests and better able to assimilate local cultural norms dictating levels and domains of competitiveness. Beginning in early childhood, implicit dominance hierarchies become established based on individual differences in physical prowess and social skills. Dominance is defined as access to resources, territory, and power over others. By middle childhood, skills that facilitate dominance expand to incorporate achievements in diverse domains unique to humans, such as intellectual, musical, artistic, or athletic performance. Individuals also become increasingly conscious of their relative positions within the hierarchy.

In adolescence, competition heightens as individuals begin to compete for reproductively attractive mates. Males compete with other males for physical and symbolic status, which attracts females, because high status can be converted into resources to support females and children. Females compete with other females on physical attractiveness and sociability, which attract potential stable marital partners. Coalitions form that permit individuals to increase jointly their positions in the hierarchy.

An established dominance hierarchy reduces overall competition within the social group and confines competition to individuals close in status. High- versus low-status positions in a hierarchy provide individuals with varying expectations and levels of confidence regarding success in achieving personal goals. Varying the social and physical environment modifies relative dominance position and concomitant

cognitive expectations, levels of confidence, and biological substrates.

Joyce Benenson

See also Aggression; Conflict; Gender Differences; Peer Influences

Further Readings

Campbell, A. (1999). Staying alive: Evolution, culture, and women's intra-sexual aggression. *Behavioral and Brain Sciences, 22*, 203–252.

Knight, G. P., & Kagan, S. (1977). Development of prosocial and competitive behaviours in Anglo-American and Mexican-American children. *Child Development, 48*, 1385–1394.

Maccoby, E. E. (1990). *The two sexes.* Cambridge, MA: Harvard University Press.

Roy, R., & Benenson, J. F. (2002). Sex and contextual effects on children's use of interference competition. *Developmental Psychology, 38*, 306–312.

CONDUCT DISORDERS

Conduct disorders are characterized as the more severe, chronic, and pervasive forms of behavioral problems. They include a range of inappropriate, destructive, and sometimes malicious actions that constitute psychopathology. There are several clinically diagnosed mental health syndromes among conduct disorders. These are based on the constellation of maladaptive personality traits exhibited. This chapter will review the more common forms: oppositional defiant disorder, conduct disorder, and antisocial personality disorder. In addition, an overview of risk factors and intervention methods will be discussed.

Conduct disorders have the potential to significantly diminish the quality of life and long-term achievement for individuals with these behaviors. In addition, there are substantial emotional and financial costs to educators, families, and communities resulting from aggressive and destructive acts. Fortunately, a variety of interventions are available that can improve prognosis, especially when applied early, before behaviors become habitual. A number of contributing risk factors have been identified; however, much research is still needed to better understand causal agents. Multifaceted intervention plans that target a variety of risk factors are noted to be

most successful. Additional research on treatment efficacy is also needed.

Understanding and Identifying Conduct Disorders

An understanding of conduct disorders is important to researchers, educators, and policymakers in several ways. Researchers require knowledge of conduct disorders in order to investigate diagnostic measures and treatment efficacies that improve educational outcomes for these students. Teachers need to understand classroom behavioral management strategies for conduct disorders to assist students with their academic achievement. School administrators must also understand behavior needs to ensure appropriate teacher training, empirically based service models, and systemic structures that support positive learning environments. Policymakers in government agencies and business provide a crucial role in designing curriculum materials, establishing regulations, and providing infrastructure for school systems. Therefore, policymakers also require an understanding of risk factors and treatments for students with conduct disorders.

Formulating decisions regarding the identification or treatment of maladaptive behavior requires appropriate expertise. The clinical skills of diagnosis can appear deceptively succinct, even simplistic, if symptoms are treated as a checklist of manifest items. However, diagnosis is a multifaceted and complex process, requiring substantial professional training. Although the clinician is provided a list of symptoms to identify, this information must be juxtaposed with a broader and more sophisticated understanding of normal development. Understanding the normal developmental trajectory of behavioral stages provides a good comparison benchmark for identifying abnormal behaviors. As they mature, typically developing children follow a predictable sequence when acquiring skills for self-regulation of their emotions and actions, stronger impulse inhibition, and self-awareness. A professional mental health provider will know the ages at which children and youth typically exhibit behaviors, even negative ones such as biting or rebellion. The same behaviors may be considered maladaptive for one child and not another based on age-appropriateness. For example, a 2-year-old child who threatens another child on the playground to secure a toy is viewed differently from an adolescent exhibiting the same behavior. Frequency

and chronicity of behaviors also are hallmarks. Criteria for conduct disorders will often require a specific number of incidents of a particular type over a period of time for diagnosis.

A number of additional issues warrant consideration in understanding the function of behaviors and conduct disorders. Clinicians must establish the level of impairment resulting from the conduct disorder. This can range from someone who is in imminent danger of harming others to no current impairment. The impairment range will help determine the level of services required for the individual, and often the Global Assessment of Functioning Scale codes are used to document this. Consideration must also be given for possible comorbid disorders and differentiating between disorders with similar features.

Knowledge of the dominant culture of the individual also may provide contextual information for behaviors that provides additional insight. For recent immigrants, societal norms and expectations may be considerably different, especially for persons escaping war or violent circumstances. A history of aggression only for self-defense would not constitute a conduct disorder. Even within the U.S. mainstream culture, there are urban settings where gang activity in neighborhoods can result in different adaptive behavioral expectations. For persons subject to these circumstances, a behavioral bravado or highly assertive stance may have adaptive value as it can lessen vulnerability to victimization.

A number of interpersonal situations also can create temporary states of inappropriate behaviors that are not pathological. Adults and children who experience multiple life stressors such as divorce, death, or trauma can exhibit negative behavioral changes. These can manifest as irritability, agitation, anger, or negative moods similar to conduct disorder features. However, negative behaviors usually subside quickly. Because of the sudden onset and temporary nature, these symptoms may be more consistent with adjustment disorders rather than conduct disorders. Some medication side effects can also mimic behavioral symptoms.

Oppositional Defiant Disorder

The *Diagnostic and Statistical Manual of Mental Disorders* (Fourth Edition, Text Revision) *(DSM–IV–TR)*, published by the American Psychiatric Association, provides the symptom criteria professionals use for mental health diagnoses. Oppositional defiant disorder

(ODD) presents as a pattern of antagonistic and defiant behaviors. Youth may exhibit a variety of angry, argumentative, and vengeful behaviors. Although symptoms can be verbally aggressive and hurtful to others, ODD is not characterized by physical violence or criminal acts. Persons with ODD often lack self-evaluative insight or acceptance of responsibility and will blame others for their own inappropriate behaviors.

Conduct Disorder

The *DSM–IV–TR* diagnosis of conduct disorder (CD) differs from ODD in that behaviors violate the rights of others and generally accepted rules. Behaviors may include physical aggression or criminal acts (e.g., burglary, vandalism, rape, arson). The perpetrators can be cunning, purposeful, and deceptive in their attempts to manipulate others. When considering long-term outcomes, youth who first exhibit CD in adolescence, rather than early childhood, have a better prognosis.

Antisocial Personality Disorder

Based on criteria in the *DSM–IV–TR*, if a pattern of chronic disregard for the rights of others and compliance with the law persists at age 18, a diagnosis of antisocial personality disorder (APD) may be appropriate. The diagnosis requires CD symptomology since age 15. Behavioral features of this diagnosis include lacking empathy or guilt when harming others, manipulative interactions with others, gross irresponsibility, and criminal activity.

Risk Factors and Potential Causal Pathways

Within the literature, a number of potential risk factors for conduct disorders have been identified. They are thought to represent an interaction of multiple factors, including vulnerability, predispositions, environmental elements, and poor personal choices, that increase the risk for severe behavior problems. These include dysfunctional communication patterns, negative familial behavioral patterns, difficult temperaments, and poor emotional self-regulation. Additional factors include poverty, deviant peer associations, lower intelligence (especially verbal), presence of multiple sociocultural stressors, and comorbidity for other mental health diagnoses.

Emotional Self-Regulation

Persons with conduct disorders often engage in high-risk behaviors without consideration for consequences or the effects on others. They can be highly reactive to emotional situations or perceived injustices. They may also report impulsive actions that happen with little premeditation. Based on these behavioral patterns, some researchers suggest that conduct disorders may be the result of poor emotional self-regulation. As children mature, they typically learn to self-monitor their own feelings, inhibit first responses, and select reactions. However, youth with conduct disorders do not display these qualities. Therefore, emotional self-control is hypothesized to be either delayed or poorly developed. In addition, a significant number of persons with conduct disorders are also diagnosed with attention deficit hyperactivity disorder (ADHD). Diagnostic features can include notable impulsivity and hyperactivity as found in ADHD combined type.

Dysfunctional Communication

There is evidence through the work of researchers such as George Patterson that dysfunctional communication patterns may contribute to some behaviors in conduct disorders. Parents who exhibit an authoritarian discipline style that is coercive in nature can escalate hostility. The cycle can be perpetuated if parents respond to or make demands in a threatening manner, the child then becomes defiant, and the parent in turn raises the level of threat or anger in the commands. This process serves to increase the potential for physical altercations as well as teach coercive interaction patterns. Other researchers, including Kenneth Dodge, have identified the propensity of students with conduct disorders to misinterpret social cues that are ambiguous. These cognitive distortions can lead to inappropriate hostile interpretations of interactions with others.

Negative Familial Behavior Patterns

As noted with many other mental health diagnoses, conduct disorders are more frequent when there is a family history of similar disorders. This suggests possible biological vulnerabilities as well as a number of social learning and environmental hypotheses. Parents who also have a history of conduct disorders or are experiencing other mental health issues (e.g., depression, schizophrenia) are likely to model behaviors that reinforce negative interaction patterns. There is evidence that mothers with severe depression may not form critical early attachments with their children. In the case of parents with APD, their daily life skills can be significantly diminished. This can result in a lack of basic parenting skills that can lead to sickness, neglect, or abuse of their children. The marital relationship and financial stability are often jeopardized because parents' irresponsible and callous actions create a less stable home environment for children. It is also likely parents with APD will have a history of prosecution and perhaps incarceration that further diminishes home stability for the child.

Temperament

Temperament traits are described as relatively stable personality characteristics that influence behaviors in a predictable manner. Early theory by Carl Jung explained temperament in terms of dichotomies such as the widely acknowledge premise of extroversion and introversion. He described two preferences for learning new information, sensing with an emphasis on facts and intuition with an emphasis on concepts. Jung also noted that when his patients had a strong preference for one temperament dimension, the opposing dimension could be underdeveloped. He noted that the pathology of his psychiatric patients was often expressed in ways consistent with their temperament. For example, extroverted patients were more likely to act out or show aggression.

In studies of students with ODD and CD by Diana Joyce and Thomas Oakland, preferences have been noted for extroversion, sensing, and thinking functions. Of particular note, the students with ODD were significantly more likely than the general population or students with CD to prefer new information presented in sensing styles. This includes explicit, direct instruction presented in sequential steps that utilize concrete, applied examples and detailed facts. These characteristics may explain the propensity of children with ODD to argue about rules and not generalizing behavioral expectations across circumstances.

In their work with very young children using a physiological-based temperament model, Alexander Thomas and Stella Chess described the concept of goodness of fit. This concept postulates that when parent and child temperament styles are dramatically different, there can be less understanding and tolerance for the child's behaviors. This lack of compatible

behavioral patterns has the potential to increase conflict. Their research found that children with difficult or irritable temperaments had poorer long-term outcomes.

Educational Institution Services

Public schools from kindergarten through twelfth grade provide a wide range of general and special education services, programs, and accommodations for children with special needs. Historically, referrals for these types of services have been made by teachers, although parents may also request evaluations. Services were primarily accessed following an evaluation and classification within the exceptional student education (ESE) system. If the student qualified, conduct disorders were often served under an Emotional Disturbance classification. Students could receive services ranging from in-class consultation to small group pull-out sessions or self-contained classrooms. In severe cases, students may be served through inpatient hospital or residential treatment programs.

Based on U.S. Department of Education statistics for 2002, the Emotionally Disturbed category serves the fourth largest number of special needs students. The criteria for ESE categories are established in the Individuals with Disabilities Education Act of 2004 and do not require specific mental health diagnoses. State departments of education also provide more narrow statutes that describe specific evaluation components. If schools are compliant with federal mandates, additional funding is provided for these services.

More recently, the provision of services within school systems has changed to include a response-to-intervention (RtI) model. This model is based on several premises. First, an emphasis on prevention and early intervention is required. Second, in-class and small group services may be accessed prior to a full evaluation or placement as Emotionally Disturbed. The goal is to address student difficulties before they become chronic or more complex. Third, each level of student service requires monitoring and review of outcome data to determine the efficacy of the intervention.

The concept of RtI is still evolving, and predominant models conceptualize three or four tiers. The first tier, as applied to behavioral disorders, includes school-wide screening for behavioral or socioemotional risk factors. This is a procedure that has not been commonly practiced in the past and may identify student needs earlier as well as some students who were not noticed by teachers in the past. The second tier is often in-class or small group intervention with an emphasis on collecting data to demonstrate the effectiveness of the remediation strategies. If significant improvement in behaviors is not noted, the final tier involves a more comprehensive psychoeducational evaluation. Based on results, school personnel may advocate more intensive, frequent, and multimodal services. With the advent of RtI, a new set of skills that includes single-subject design, treatment protocol studies, and systemic personnel curriculum training will be required of educators.

Intervention Strategies

Numerous negative outcomes have been associated with conduct disorders. These include considerable financial costs to schools and communities as well as emotional distress for individuals. Students with conduct disorders receiving services for Emotional Disturbance require additional funding support and school personnel resources. They experience more discipline referrals that consume the time of behavioral deans, in-school suspension personnel, other support personnel, and expulsion proceedings. In some cases, considerable additional cost is incurred to contract with outside agencies for specialized services such as counseling therapy. Data indicate higher school failure, higher dropout rates, lower incomes, and more frequent incarceration. These factors bear substantial emotional and mental health cost for individuals with conduct disorders who do not meet their full academic potential. In lieu of both the personal and systemic costs, early intervention is crucial.

When considering interventions, it is important to note that the incidence of conduct disorders is higher for males than females, especially early childhood onset. This is an important distinction because early onset of conduct disorders is associated with poorer outcomes. There also are gender differences in symptoms, suggesting that differing interventions may be required. Girls have greater risk for prostitution, alcohol or drug use, skipping school, lying, and running away from home. Boys have greater risk for physical altercations, theft, and destruction of property. Acquiring thorough background and evaluation data for individuals can provide considerable guidance for specific needs when selecting interventions.

Socioemotional Evaluation

The initial step toward evaluation may result from school-wide screenings for risk factors or teacher/parent

recommendations. Teacher and parent clinical interviews are conducted to determine the areas of concern. The goals of interviews include documenting medical and educational history, gathering pertinent familial history, recording a temporal sequence of prior events, establishing age of symptom onset, and noting any extenuating circumstances. Topics are guided by the clinician and may include issues of family dynamics, peer relationships, and juvenile delinquency incidents. Prior remediation strategies used for the student and their effectiveness, as well as student strengths, are solicited to inform later intervention choices. Consideration for student strengths and interests is particularly important to fostering early successes during intervention and identifying reinforcers for behavioral modification plans. Behaviors of concern are noted in clear, objective descriptions that can be measured later.

With objective and measurable descriptions of target behaviors, direct observations can be conducted. These data have many formats, including anecdotal records that indicate antecedents and consequences, as well as frequency and duration measures. Data collected prior to intervention can serve as the baseline for treatment. They may also be analyzed to indicate patterns, such as time of day, particular tasks that prompt behaviors, or specific locations in which behaviors escalate.

Several nationally recognized omnibus rating scales are available that sample a broad range of externalizing and maladaptive behaviors. These scales may be administered to teachers, parents, and the students themselves. They provide a cross-informant analysis to identify patterns of concern. Norm-referenced T scores are calculated to indicate average, at-risk, and clinically significant range behaviors in comparison to other individuals their own age. Numerous single-construct measures are also available to explore more narrow issues (e.g., opposition) with a greater number of items and specificity.

Because conduct disorders are often associated with poor adaptive skills, particularly interpersonal relations and communication, an adaptive measure may be used. Adaptive rating scales measure social, conceptual, and everyday practical skills. In addition, there are temperament, social skills, study habit, and self-concept measures that will help provide a more comprehensive review of the student's skills. Intelligence and academic measures can identify general cognitive abilities and specific academic skills. Addressing academic needs is particularly important to a multimodal intervention design because learning has an impact on task frustration, school motivation, and long-term achievement.

Systemic Interventions for Conduct Disorders

Systemic interventions can be applied when school-wide changes in behavior are needed because problems affect many students. These procedures may be most useful when there are frequent behavioral discipline incidents within a particular school or several students are engaging in behaviors that victimize others routinely, such as bullying.

Positive Behavioral Support

Positive behavioral support is a systematic intervention that focuses on ensuring that the school environment provides a positive, nurturing climate that reinforces appropriate actions. The focus is on lowering behavioral referrals and preventing discipline incidents rather than reacting after misconduct has occurred. It is recommended that strategies be implemented school-wide across classrooms and grades. This provides a clear, consistent structure where students know the expectations for behaviors. Desired behaviors are stated in prosocial terms (e.g., personal responsibility, sharing, respect, cooperation) and emphasize the well-being of all students, as well as school pride. Some programs include public reinforcement or rewards that provide recognition for appropriate behaviors (e.g., good citizen prizes). Expectations for positive behaviors are clearly communicated throughout the school. This can be accomplished through posters, banners, school slogans, and even morning announcement slogans. Because these strategies are comprehensive, the participation of administrators and teachers is crucial. This can require some initial training to ensure that all school personnel understand and reinforce the concepts.

Environmental Structure Changes

Another systemic intervention within schools to control disorderly conduct is the strategy of incidence mapping with structural modification. The goal is to prevent or lower disruptive incidents by limiting the opportunities for covert actions. A review of the campus grounds is conducted with an analysis of all the

schools behavioral discipline referrals. The objective is to identify patterns of incidences related to environmental routines or structures. Examples include frequent incidents in isolated areas, unsupervised areas, or areas that are not well lit (e.g., bullying in bathrooms). Easy preventive measures to lower these incidents include better lighting, ensuring that adult supervision is nearby, and sometimes changing structures. Incidents may also cluster around certain activities, times of day, or locations that are overcrowded. Examples include busy cafeterias, too many student lines passing simultaneously in crowded halls, or rowdiness during bus loading. Easy remedies include changing travel routes, staggering access to areas, or increased supervision.

Social Learning Curriculum

A number of published curriculum packages are available to schools that incorporate prosocial themes. These may be integrated into academic curricula to reinforce concepts of altruism, self-control for emotions, good problem-solving skills, and even social skills. They may also be introduced periodically, often through school counselors, as special sessions to remind students of appropriate behaviors and reasons for those choices.

Small Group and Individual Interventions for Conduct Disorders

A number of small group and individual interventions have demonstrated positive effects for diminishing conduct problems. Small group interventions are often limited to a few weeks and target a specific common need that several students may have. Examples include social skills training, anger management, conflict resolution, grief counseling, or test anxiety/frustration. A number of small group curricula are available, and they generally have the following components: opening exercises for rapport, identification of feelings, teaching of three to five new coping strategies, practice to generalize the new skill, and a follow-up or booster session.

Behavioral Modification

Operant conditioning principles are the basis for behavioral modification. This technique typically targets specific, individual behavioral goals. Reinforcers

and punishment can be used to promote desired behaviors or lower inappropriate behaviors. Reinforcers may include verbal praise, tangible prizes, token economies, and the opportunity to engage in preferred tasks. Punishments may include time-out and overcorrection techniques. Once specific behaviors are targeted, baseline data on the frequency of the behavior are collected. As the behavioral modification plan is applied and appropriate behaviors are reinforced, data are collected to determine the effectiveness of the intervention.

Cognitive Behavioral Therapy

Cognitive behavioral therapy is based on the premise that distorted or dysfunctional thinking patterns can result in maladaptive behaviors. Therefore, the target of intervention is to identify and change existing thinking patterns to new, more adaptive patterns. Methods include teaching relaxation training, gradual exposure to a feared situation, learning replacement thoughts, and biofeedback to control physiological reactions. Applications to conduct disorders include teaching students to recognize the precursors to explosive anger, providing self-calming strategies to lower arousal states, and modeling alternative problem-solving solutions.

Parent Training

Parent training can be an effective intervention for conduct disorders. In some cases, school personnel and parents may collaborate to provide a consistent behavioral modification plan across the two settings. Parenting programs often stress noticing the positive behaviors of children, using reinforcement to shape behaviors, and providing clear expectations for behavior. Authors such as Russell Barkley have designed training modules that teach parents of defiant children how to implement successful behavioral strategies at home. Techniques include the use of a token economy and time-out methods.

Psychopharmacology

There are a number of medications used to treat behavioral symptoms for some persons with conduct disorders. Stimulants may be prescribed when conduct problems are related to ADHD, impulsiveness, or hyperactivity. These include Adderall, Concerta, Cylert,

Dexedrine, and Ritalin. A class of medications called selective serotonin reuptake inhibitors (SSRIs) is often used to treat depression and panic syndromes. Paxil, Prozac, and Zoloft are common SSRIs. Wellbutrin, Effexor, Risperdal, and Depakote may lessen behavioral symptoms related to anxiety, severe mood changes, and bipolar disorder. The use of medications for children is controversial because limited research studies demonstrating treatment efficacy are available for some age groups. In addition, there are few studies on the possible long-term side effects of drug usage for children.

Diana Joyce

See also Aggression; Attention Deficit Hyperactivity Disorder; Behavior Disorders; Behavior Modification; Bullying

Further Readings

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., Text rev.). Washington, DC: Author.

Barkley, R. A. (1997). *Defiant children: A clinician's manual for parent training.* New York: Guilford.

Barkley, R. A., & Mash, E. J. (2003). *Child psychopathology* (2nd ed.). New York: Guilford.

Dodge, K. A. (1993). The future of research on the treatment of conduct disorder. *Development and Psychopathology, 5,* 311–320.

Frick, P. J. (1998). *Conduct disorders and severe antisocial behavior.* New York: Plenum.

Individuals with Disabilities Education Improvement Act of 2004. 20 U.S.C. § 1400 et seq. Retrieved November 2, 2006, from http://www.idea.ed.gov

Joyce, D., & Oakland, T. (2005). Temperament differences among children with conduct disorder and oppositional defiant disorder. *California School Psychologist, 10,* 125–136.

Jung, C. G. (1971). *Psychological types* (R. F. C. Hull, revision of trans. by H. G. Baynes). Princeton, NJ: Princeton University Press. (Original work published 1921)

Patterson, G. R. (1982). *Coercive family process.* Eugene, OR: Castalia.

Thomas, A., & Chess, S. (1989). Temperament and personality. In G. A. Kohnstamm, J. E. Bates, & M. K. Rothbart (Eds.), *Temperament in childhood* (pp. 249–261). New York: Wiley.

U.S. Department of Education. (2002). Twenty-fourth annual report to Congress on the implementation of the Individuals with Disabilities Education Act. Jessup, MD: Author.

CONFIDENCE INTERVAL

Quantitative research in educational psychology benefits from using sample data to describe the wider population of interest. A confidence interval (CI) is a statistical approach that uses sample data to determine a range of values on a particular statistic (e.g., mean) for a sample from the population. In theory, a confidence interval is based on an infinite number of random samples from the population. For each of these random samples, a statistic of interest, such as the mean, could be calculated. Of course, it is highly unlikely that each of these means will be equal to one another, and even more unlikely that any or all of the means will equal the population mean. Thus, it is generally safe to assume that error will be present when estimating a population parameter with a statistic based on a random sample. Comparing the CI to the sample mean provides the researcher a means for determining the magnitude of error.

A CI consists of two numbers, often referred to as lower and upper bounds, between which the value of a population parameter is expected to fall. For instance, the CI for a population mean is typically reported in the following format:

$$CI_{PI} = LowerBound < \mu < UpperBound,$$

where CI_{PI} indicates the probability factor for the respective confidence interval, and μ represents the population mean. In general, probability factors of 90, 95, and 99 are used to construct CIs. For interpretation purposes, these probability factors are converted to percentages. For instance, using a probability factor of 95 results in a 95% CI, which indicates that the researcher can be 95% confident that the population mean falls within the lower and upper bounds of the interval. As another example, consider a 95% CI for mean scores on a standardized mathematics assessment for seventh graders at Jayhawk Middle School:

$$CI_{95} = 70 < \mu < 90.$$

This CI indicates that 95% of random samples taken from seventh graders at Jayhawk Middle School will yield a mean math score greater than 70 and less than 90.

The latter example can also be used to illustrate the role of measurement error in estimating population parameters. Begin by assuming that the CI for seventh graders at Jayhawk Middle School was based on a sample mean that equals the population mean of 80. Furthermore, suppose a random sample of seventh graders from Jayhawk Middle School yielded a mean mathematics assessment score of 81. An additional, independent, random sample taken from the same population of seventh graders at Jayhawk Middle School yielded a mean mathematics assessment score of 77. In both cases, the sample means do not equal the population mean, nor do they equal one another.

The amount of measurement error is used to construct a CI around a particular parameter. This measurement error is referred to as the standard error. The standard error provides a measure of the variability of the sample statistic around the population parameter. For instance, the standard error of the mean is calculated using the following formula:

$$\sigma_{\bar{X}} = \frac{s}{\sqrt{N}},$$

where σ_X is the standard error of the mean, s is the standard deviation of the sample, and N is the sample size. The distribution of an infinite number of randomly sampled means is known to be normal and can be converted to a standard normal distribution. Values can be taken from the standard normal distribution to construct CIs. For instance, a 95% CI requires the use of a standard value of 1.96. This value is multiplied by the standard error of the mean in order to convert the CI to the original unit of measurement. Therefore, the equation for constructing a 95% confidence interval is

$$\bar{X} - 1.96\sigma_{\bar{X}} \leq \mu \leq \bar{X} + 1.96\sigma_{\bar{X}}.$$

As seen in the equation, the product of the standard value and the standard error are subtracted and added to the sample mean to create the lower and upper bounds for the CI.

CIs can also be used to test research hypotheses. The most straightforward hypotheses involve determining if a sample mean is representative of the population. The hypothesized value of the population mean is considered tenable if it falls within the bounds of the CI created by the randomly sampled mean. However, if the hypothesized population mean

does not fall in the CI, the researcher has obtained evidence that the population mean is different from the hypothesized value.

To illustrate this concept, reconsider the example using students from Jayhawk Middle School, where two random samples yielded sample means of 81 and 77, respectively. In addition, suppose a researcher hypothesizes that the population mean score for seventh-grade students from Jayhawk Middle School is 80. For simplicity, assume $s = 2$ and $N = 4$ for both samples, which results in a standard error of 1 for both samples. Using this information, a 95% CI can be constructed for each sample. The 95% CI for Sample 1 and Sample 2 are, respectively,

$$79.04 < \mu < 82.96$$
$$75.04 < \mu < 78.96.$$

The CIs result in different interpretations regarding the hypothesis. In Sample 1, the hypothesized value is contained in the CI, indicating that a mean score of 80 for seventh-grade students from Jayhawk Middle School is tenable. On the other hand, the CI constructed from Sample 2 data does not contain the hypothesized value of 80. Therefore, based on this sample, the mean score for seventh-grade students from Jayhawk Middle School is not equal to 80. In fact, the researcher could state with 95% confidence that the mean score instead falls between 75.04 and 78.96.

In summary, CIs are one of the most prevalent forms of interval estimation used in the behavioral sciences. The width of a confidence interval can give the researcher an idea about the uncertainty in the precision of the parameter that he or she is trying to estimate. For example, a "wide" CI indicates that there is error in the measurement and further data collection should be employed. In addition, CIs are considered more informative than hypothesis tests because they provide a probable range for the unknown parameter.

Greg W. Welch and Chris S. Meiers

See also Inferential Statistics; Scientific Method; Statistical Significance

Further Readings

Smithson, M. (2003). *Confidence intervals.* Sage University Paper series on Quantitative Applications in the Social Sciences, 07–140. Thousand Oaks, CA: Sage.

CONFLICT

Conflict is a disagreement between two or more inter-dependent parties about goals and/or methods for achieving goals. Although some scholars use the term *conflict* to refer to an internal tension or confusion (for example, a mother has "role conflict" when she is torn between competing demands on her time and attention due to her role as a parent and as a working professional), conflict is a social phenomenon concerning disagreements between parties.

A party to a conflict can be a person, a group, an organization, a state, or a nation. Often, conflicts are between two parties who are individuals. But conflict can involve many parties. The more parties to a conflict, the more difficult it is to manage the conflict process and to come to a decision or settlement that is acceptable to all parties.

All parties to a conflict are interdependent; they depend on each other for some action or resource necessary to manage or resolve the conflict. The more one party is needed by the other parties, the more power and influence that party has to define the conflict, determine when and how the conflict is managed, and influence the final outcome or solution for the conflict.

What types of disagreements lead to conflict? Conflicts may be caused by disagreements over scarce resources and how they should be distributed. Scarce resources may be economic, such as money, equipment, and capital; or they can be relational resources, such as status, respect, information, and time. Sometimes, conflicts are disagreements about perceptions that one party is exercising unwanted control over other parties. Conflicts are often disagreements about goals—questions about what should be accomplished and why. Even if parties agree on goals or desired outcomes, they may disagree on the best methods to achieve the outcomes. For example, office managers may agree that they need to upgrade their computer systems, but they disagree about which computer system is the best for their organization.

Conflict can be functional or dysfunctional depending on how the conflict is managed. In functional conflicts, parties are satisfied with both the process used to manage the conflict and the resolution to the conflict. Learning functional conflict management is important for social development and relational development. From early childhood, people learn how to be in peaceful relationships by understanding how to manage conflicts constructively. Many developmental psychologists identify learning to manage conflict constructively as an important skill, and perhaps the most important social skill, of childhood. Educators who work to develop children's emotional intelligence and emotional competence also see constructive conflict management as a key skill for the emotionally competent person. When a conflict is functional, the parties are open and willing to honestly share information, there is a sense of calm and respect, parties are cooperating to try to solve the problem together, and they are flexible in looking at options for solutions.

Dysfunctional conflict happens when a party is dissatisfied with the process or outcome of the conflict. For example, one party may be upset that the other party is unwilling to go to a third party (for example, a mediator) for help with the conflict, or feels the solution to the conflict is unfair. When conflicts become dysfunctional, they escalate in terms of increasingly negative behaviors. Dysfunctional conflicts are characterized by distrust, unwillingness to share information openly and honestly, tension, and the adoption of an "us versus them" attitude that leads to competitive and aggressive behavior. Dysfunctional conflicts often lose focus, and the parties start arguing about issues other than those that initially triggered the conflict. Consider a husband and wife who begin arguing about money and are suddenly arguing about in-laws, parenting, and household chores as well. When conflict becomes dysfunctional, it affects future conflicts. If a person feels he or she was treated unfairly, the person is likely to be defensive and aggressive in the next conflict with the same party.

There are five ways to manage conflicts without the aid of a third party: competition, cooperation, compromise, accommodation, and avoidance. If used in the right situation, all of these techniques or styles can lead to functional conflict. Competition is a win-lose approach, and the goal is to get as much for your side as possible. Competition does not have to be violent or aggressive; for example, buying a car is usually done through bargaining, which is a competitive approach to managing conflict. Cooperation, also called collaboration, is a win-win approach in which parties problem solve and generate solutions that meet everyone's needs. Cooperation is best when there are creative approaches to maximizing gain and when parties must be committed to implementing solutions. Compromise is a give-and-take approach that works well when there are limited resources. However, resources may not be that limited, and cooperation may be the better

approach. Accommodation is letting the other party get what he or she wants in the conflict; this strategy is effective if relational harmony is most important. The danger is establishing expectations that one will usually accommodate. Avoidance is physically and/or psychologically withdrawing from the conflict. For unimportant or short-term conflict, avoidance is a feasible option, but it is counterproductive when the conflict will become worse if not addressed.

Conflict management can also depend on third-party processes such as mediation, arbitration, and adjudication to solve the conflict. Conflict resolution education programs help teach youth and adults constructive ways of managing conflict.

Tricia S. Jones

See also Aggression; Competition; Cooperative Learning; Emotional Development; Emotional Intelligence; Empathy; Psychosocial Development; Social Development

Further Readings

Canary, D. J., Cupach, W. R., & Messman, S. J. (1995). *Relationship conflict.* Thousand Oaks, CA: Sage.

Deutsch, M. (1973). *The resolution of conflict.* New Haven, CT: Yale University Press.

Deutsch, M., Coleman, P. T., & Marcus, E. C. (Eds.). (2006). *The handbook of conflict resolution: Theory and practice* (2nd ed.). San Francisco: Jossey-Bass.

Hocker, J. L., & Wilmot, W. W. (2003). *Interpersonal conflict* (6th ed.). Dubuque, IA: William C. Brown.

Pearce, W. B., & Littlejohn, S. W. (1997). *Moral conflict: When social worlds collide.* Thousand Oaks, CA: Sage.

CONSERVATION

Conservation refers to the feature of constancy or invariance of an object. It is the notion that a quantity stays the same despite changes in form or appearance. Educational psychologists study the concept of conservation in children as an important developmental milestone that is reached when a shift in cognitive thinking occurs. One of Jean Piaget's earliest and most important discoveries was that young children lacked the principle of conservation. Much research has been conducted to prove or disprove Piaget's theory of conservation.

While studying how children develop intelligence, Piaget discovered that young children do not understand that quantity, length, or number of items is unrelated to the arrangement or appearance of the items. He developed different tasks to further understand the concept of how children develop the principle of conservation. He looked at mass, number, volume, and area.

For the task of mass, the examiner presents a child with two round balls of clay equal in weight, size, and shape. The child is asked if the two balls are the same weight. The child is allowed to use a balance to assure the weight is equal if needed. The examiner then transforms one ball of clay into a flat pancake or a thin rope. The child is asked again if the two quantities of clay have the same weight and to give and explanation of why or why not.

For the task of number, the examiner presents the child with two equal rows of counters. The rows of counters are arranged side by side in one-to-one correspondence. The child is asked if the two rows are equal. The examiner then spreads one row of coins apart, and again asks the child if the two rows have equal amounts of coins.

The task to assess conservation of volume begins with showing the child two identical containers filled with the same amounts of liquid and asking the child if the amount of liquid is equal. Then, the liquid from one container is poured into a container that is taller and thinner than the other. The child is again asked if the amount of liquid is the same in both containers.

A fourth, less used, and less researched task used by Piaget was the "cows on a farm" test for conservation of area. In this task, two identical farms were presented, each with a cow placed on it. The children were asked if each cow had the same amount of grass to eat. Then, Piaget would add little cubic farmhouses to each farm. On one farm he would line up the farmhouses in a tidy row, on the other farm, he would scatter the farms about on the grass. The child is then asked again if the cows have the same amount of grass to eat.

In each of the above tasks, the child's ability to recognize that certain physical properties do not change, or are conserved, even though they may undergo a physical transformation, is tested. Children who do not understand the principle of conservation will say that the quantity of mass, number, volume, or area has changed.

The ability to recognize the principle of conservation is the dividing line between Piaget's preoperational stage and concrete operations stage. Most

children move from one stage to the next between the ages of 6 and 8.

Piaget believed that children who were unable to conserve mistook a perceptual change for a quantitative change. This confusion could arise from the pre-operational child's tendency to focus or center on only one characteristic of an object to the exclusion of all other characteristics. Additionally, the child's inability to mentally reverse actions is characteristic of pre-operational thought. For example, the pre-operational child cannot mentally reverse the clay-rolling process to determine if the balls of clay are still the same size. A child in the concrete operational stage can understand that nothing has been added or subtracted to the ball of clay. This child can understand that the two dimensions of height and weight are related. The development from pre-operational to concrete operational is a gradual one. A child may be able to achieve number conservation before liquid or mass conservation.

Research regarding Piaget's theory of the principle of conservation has presented information that children may have difficulty with the discourse of the interview that accompanies the conservation tasks. These studies indicate that children may display an understanding of conservation principles earlier than previously thought but misunderstand what is being asked of them by the experimenter. Additionally, some research has shown that with minimal training, competence for the tasks increases.

Other research has pointed that Piaget's tasks are, in fact, testing two different forms of conservation: identity and equivalence. The conservation of identity must be inferred from the child's responses, and the conservation of equivalence can be evidenced in the child's judgments. This means that a child is actually concerned with identity conservation, or how the item looks different but is still the same, but in reality, the child is making an equivalence judgment. But children at this age do not verbalize the actual psychological processes involved in their judgments; instead, they apply logical reasoning.

Conservation is learned through experiences and interactions. The process is gradual and involves developing the ability to focus on more than one attribute of an object at one time, the understanding of reversibility, and the ability to mentally represent objects.

Lori Jackson

See also Cognitive View of Learning; Learning Style; Piaget's Theory of Cognitive Development

Further Readings

Elkind, D. (1967). Piaget's conservation problems. *Child Development*, 38, 15–28.

Piaget, J. (1963). *The origins of intelligence in children.* New York: Norton.

Siegal, M. (1991). A clash of conversational worlds: Interpreting cognitive development through communication. In J. M. Levine & L. B. Resnick (Eds.), *Socially shared cognition* (pp. 23–40). Washington, DC: American Psychological Association.

CONSTRUCTIVISM

Constructivism is a learning theory based on the notion that students actively construct knowledge. This view of learning calls for a dramatic lessening of reliance on a didactic, textbook-based, "transmission" of knowledge approach to teaching and learning in the classroom. It also challenges popular views about the nature of knowledge. The fact that constructivism rejects commonplace views about knowledge, however, does not mean that it embraces a relativistic or anything-goes approach. What is also clear is that constructivists who endorse views that oppose the mainstream view—who argue, in other words, that worthwhile knowledge tends to be complex rather than simple, open to question rather than certain—are more likely than acknowledged to promote disagreement among experts about how best to conceptualize knowledge.

Research shows that these core beliefs about knowledge relate to teacher practice, albeit imperfectly, because of the pressures teachers face to cover the curriculum and improve standardized test performance. Another factor that affects teachers as they seek to convert theories of knowledge into classroom practice is that constructivists who argue for a more open approach to teaching and learning disagree about the particulars of the openness. Some constructivists, for example, maintain that the curriculum should be problem based; others argue that it should be based on the students' interest and level of development. The same disagreement arises about the nature of the pedagogy. Some constructivists believe that it should be peer centered; others maintain that it should be individually oriented. A similar disagreement exists with regard to assessment. Some argue that it should be performance based, which is to say, out in the open and judged against a common standard; others believe

that it should be based primarily on each child's progress, best assessed using individual portfolios.

These disagreements about practice among constructivists mirror other, sometimes subtle, theoretical disagreements between different "camps" of constructivists. The disagreements turn largely on the question of what kind of knowledge is most important and how that knowledge is best acquired. Although this may be an oversimplification, one camp of constructivists, the social constructivists, want to move knowledge out of the head and into the open. In this view, knowledge is a community and not an individual possession, a notion that has implications for how members of this constructivist camp define the nature of knowledge. For one, if it really is between people, it has to be overt or observable, which means that knowledge can take two possible forms: It can consist of strategies or routines, or, favored by those social constructivists who consider themselves "postmodern," knowledge can be defined as language.

The two groups of social constructivists, because they define knowledge in two different ways, also hold up different models of learning as being ideal for the classroom. The group that focuses on strategies, strongly influenced by a Russian psychologist named Lev Vygotsky, advances apprenticeship learning as the ideal. The intellectual version of apprenticeship learning can involve a relatively simple strategy such as a mnemonic code (e.g., "one-bun, two-shoe …"), which enables one to visualize an ordered set of objects; or it can involve a complex strategy that encompasses "metacognition" (e.g., planning, monitoring, checking, and revising). Like a veteran tailor, the teacher using this approach is expected to first model the practice, gradually turning over control to the novice as he or she becomes more skillful in implementing the practice.

The second group of social constructivists views knowledge as language, as modes of discourse acquired by participating in what might be considered different types of disciplinary "language games." Language, in this approach, does not hold knowledge; rather, it helps manage or coordinate different kinds of cultural activity (e.g., "doing math"). There is both a process and a content aspect to this coordination. An example of the former might be a teacher who, because she wants her students to discuss ideas in mathematics, models language that depersonalizes the process and thus protects the self (e.g., learns to say, "It is possible that …" vs. "I think that …"). An

example of talking content at the elementary level would be students who, when they are engaged in solving math problems, consistently describe numbers as consisting of units of 10.

The third group of constructivists traces its approach back to John Dewey and, more recently, the Swiss psychologist Jean Piaget. The basic notion here is that children, and adults for that matter, are naturally driven to order or make more coherent the diverse, often fragmentary experiences they encounter each and every minute of their lives. Piaget believed that children develop increasingly complex structures or "operations" over time that allow them to create this coherence. One of the early operations, typically evident in 6-year-olds, allows them to conserve quantity—to understand, for example, that when one compares the amount of liquid in two containers, height can be offset by greater dispersion so that, although the quantity may look different, it may actually be the same. Piaget's view has fallen out of favor, replaced by Dewey's brand of inductionist constructivism. Unlike the traditional didactic approach, often characterized as "the sage (wise person) on the stage" approach, Dewey favors "the guide on the side" approach. The assumption here is that a student can create meaning only by working in his or her own experiential workspace, the 4 or 5 inches of brain between the ears. The role of the teacher is to quietly nudge the process along, to point out in a gentle way any problems the student may be encountering in figuring out how to construe a new experience, to bring to the fore the most important aspects of that experience, and so forth. The type of pedagogy that best fits this view of learning is problem-based, project-based, or inquiry-based pedagogy.

Richard S. Prawat

See also Cognitive View of Learning; Piaget's Theory of Cognitive Development; Teaching Strategies

Further Readings

Moyles, J. R., & Robinson, G. (Eds.). (2002). *Beginning teaching: Beginning learning.* Berkshire, UK: Open University Press.

Prawat, R. S. (2003). Variants on a common theme? Researching the philosophical roots of our current epistemologies. *Issues in Education, 8,* 205–215.

Wood, D. (1998). *How children think and learn.* Oxford, UK: Blackwell.

Contingency Contracts

A contingency contract, or "behavioral" contract, is a clearly written document specifying available rewards or consequences that are contingent upon the performance of a certain behavior. The use of a contingency contract is a behavioral management technique that is effective in modifying a multitude of behaviors performed in a variety of settings. A contract is based on simple behavioral principles but provides a complex intervention package. Implementing the key components and procedures outlined in this entry will enhance the effectiveness of a contingency contract.

Contingency contracts have been used in a variety of situations to increase the occurrence of appropriate/desired behavior, healthy behavior, or compliance. These types of contracts are effective in a variety of settings, including classrooms, schools, households, and clinic settings. Many individuals can partake in a contract; however, most contracts include two parties: one party who performs the targeted behavior and one party that grants access to the contingent reward.

Heward and Martin and Pear identify several advantages for using a contingency contract. First, developing and implementing a contract is a fairly simple procedure that is easy to learn and can be completed by non-professionals. Second, a contract provides a socially acceptable and positive behavior management strategy that is individualized for a child or a group of children. Third, a contract directly involves all participating individuals in the behavioral plan and development of the contract. Fourth, using a contract ensures that all participants agree with the behavior and contingent rewards specified in the document. Finally, the act of signing a contract is a powerful stimulus in our society that increases one's commitment to the agreement, thereby enhancing the integrity of implementation and consistency of delivering consequences.

Key Components for Effectiveness

Heward also outlines three major components that should be included in a contingency contract for it to be effective: (1) the *task*, (2) the *reward*, and (3) a *task record*. The *task* portion of the contract includes a description of the task, who is to perform the task (e.g., the role of each participant), when it is to be performed, and how well it needs to be performed in order to receive the reward. The *reward* section includes a description of the reward, who delivers the reward, when it will be delivered, and how much of it will be delivered. A *task record* provides a location on the document where the progress of a contract is recorded. This section for data collection includes records of task completion and reward delivery. Monitoring the progress of a contract serves two purposes: (1) it ensures that all participants are regularly reminded of the contract, and (2) it helps remind the student to continue working toward task completion.

Key Procedures for Effectiveness

A contingency contract is a complex intervention package consisting of both positive and negative reinforcement. Contingency contracts are based on rule-governed behavior, where a specific behavior results in a specific consequence. Heward and Martin and Pear have identified several procedures that make contracts more effective: (a) All participants should have equal parts in the development of the contract. (b) The contract needs to be as clear as possible so there is no confusion about what is required from both parties. (c) The target behavior needs to be clearly defined, observable, and recordable. (d) The parties should identify possible rewards that will motivate the child. (e) Data collection procedures used to evaluate the effectiveness of the contract should be determined, including how, when, and who will collect the data. (f) The relationship between the behavior and the reward should be fair. (g) Each participant must follow the rules set up by the contract. (h) The contract needs to be honest; all participants will comply with the agreement. (i) The contract should be displayed publicly, reminding each participant of the agreement and his or her responsibilities within the contract. (j) Once a clear contract is written and implemented, data should be evaluated for effectiveness. (k) The contract should be altered and rewritten until there is an improvement in behavior. It is necessary to renegotiate the contract after a certain period, even if it was successful.

Contracts in Educational Settings

Contingency contracts have been used effectively in classrooms and school settings. Typically involving a teacher–student contract, these formal agreements have been demonstrated to increase appropriate social

and academic behaviors. As an intervention, contracts between teachers and students have increased student adherence to classroom rules, compliance with teacher directions, work completion, and attendance. A contract is an extremely flexible behavioral management technique: Those established in classroom or school settings can be applied to individual students, small groups, whole classes, or entire schools. Furthermore, rewards can be provided based on individual or group-oriented contingent behavior.

Behavioral interventions using *home-school notes* are essentially contingency contracts between a child, a teacher, and a parent. They have been used to effectively modify classroom (i.e., on-task behavior) and home behavior (i.e., homework completion). There are several advantages to contingency contracts that involve individuals from both home and school. First, they allow teachers to utilize reinforcers not typically present in classrooms. Second, they enhance consistency between the two most important settings (i.e., home and school) in a child's environment. Third, they facilitate the development of home-school partnerships that are beneficial for the academic, behavioral, and social development of children.

Another type of contracting that is useful with children in educational settings is a *self-contract*. Self-contracting is a way to teach children self-management strategies. It is a self-directed method of problem solving, one that offers an opportunity for the students to take responsibility for their behavior. Included in self-contracting are the following steps: (a) self-identification of target behavior, (b) self-monitoring of behavioral performance, (c) self-evaluation of performance, and (d) self-administered consequences based on contingency.

Limitations of Contracts

Contingency contracts operate within a delayed reward framework; the reward can be perceived as far removed from the behavior targeted. Thus, contracts may be less effective when used with students who are very young or who have extremely little motivation. This is especially true when the contract requires the performance of a behavior over a period of sequential days in order to earn a reward. Other limitations include teacher and parent perceptions of using contracts. First, teachers and parents may view contracts as requiring too much time to implement or as being too complex. Additionally, they may perceive

a contingency contract as a behavioral modification technique that reduces the degree of responsibility children have for their own behavior.

John Warren Eagle and Christopher R. Niileksela

See also Applied Behavior Analysis; Behavior Modification; Operant Conditioning; Premack Principle

Further Readings

Heward, W. (1987). Contingency contracting. In J. O. Cooper, T. J. Heron, & W. J. Heward (Eds.), *Applied behavior analysis.* Columbus, OH: Merrill.

Rhode, G., Jenson, W. R., & Reavis, H. K. (1993). *The tough kid book: Practical classroom management strategies.* Longmont, CO: Sopris West.

Maag, J. W. (2004). *Behavior management: From theoretical implications to practical applications* (2nd ed.). San Diego, CA: Singular.

Martin, G., & Pear, J. (2005). *Behavior modification: What it is and how to do it* (8th ed.). Upper Saddle River, NJ: Prentice Hall.

CONTINUITY AND DISCONTINUITY IN LEARNING

Continuity and discontinuity in learning refer to the problem of whether there are qualitative changes in the fundamental mechanisms that govern learning. This question has been examined in the literature in two separate but related ways. The first approach examines whether qualitative breaks occur with increasing expertise or additional learning, and it was conducted mostly prior to the 1980s using animal learning paradigms and human learning of simple verbal materials. The second approach evaluates whether there are developmental changes in learning that are qualitative in nature.

The historical debate was inspired in part by the observation of seemingly contradictory phenomena. On one hand, much research in learning, including the seminal results from Hermann Ebbinghaus, revealed canonical smooth learning curves, in which performance could be plotted as a negatively accelerating approach to an asymptote. On the other hand, research by Robert Yerkes and Wolfgang Köhler using primates appeared to reveal "insight" in basic problem-solving tasks. Complementing these reports

were results from learning in rodents in which learning appeared to hover quite stably around chance levels of accuracy prior to rising rapidly to an asymptote. Even in tasks in which learning appears continuous, theorists such as William Estes and Randy Gallistel have shown that the learning functions for individual subjects may exhibit a more abrupt rise in performance than is comfortably accommodated within a view of learning that is purely continuous.

This dilemma motivated empirical and theoretical demonstrations from notable researchers like Kenneth Spence and Isadore Krechevsky (later David Krech) that continuous and gradual learning processes could underlie the apparently discontinuous learning curves that appeared to imply underlying learning processes that were more all-or-none than continuous. Such considerations probably played a large role in the sea-change in American psychology from stimulus-response theories to the mentalistic types of theorizing characterized by the work of Edward Tolman—theories that included previously taboo terms such as *hypothesis generation and testing* and *insight*, and that eventually became known as the cognitive revolution from the 1950s to the 1970s.

During that latter period, questions of continuity inspired one of the first widespread applications of basic mathematical models of psychological processes to questions about learning. The burgeoning set of researchers who pioneered the application of Markov processes to cognition, including William Estes, Patrick Suppes, James Greeno, and Gordon Bower, found that smooth learning curves could be produced by all-or-none learning functions, and that those all-or-none models accounted well for verbal learning data in humans. Two related empirical outcomes are central to the claim of supremacy for all-or-none models.

First, Estes showed that people tend not to recall items correctly on a second test when they have not correctly recalled them on a previous test. If discontinuous learning actually arose from variable levels of underlying learning, rather than the pure absence of learning, then an above-chance rate of recovery for previously unrecalled items would be expected. Second, there is no increase in memory performance with additional learning when only those trials prior to the final error for a particular item are examined. This, too, is consistent with the idea that those items reside in an as-yet unlearned state, rather than a state of

gradually increasing learning that is nonetheless insufficient to support accurate performance.

However, the validity of these data and of the all-or-none models was questioned, and other data appeared to contradict the central results. For example, giving subjects an opportunity to guess again when they have made an error leads to a violation of the "stationarity" principle represented by the second finding described above. The lack of a clear resolution can be best understood in light of the insights of Frank Restle, who pointed out that numerous versions of the continuous/discontinuous debate were in play, and that theorists were debating on incompatible terms. A summary of the literature prior to the mid-1960s revealed that, although it is quite clear that much learning is incremental, it is also evident that some learning is more likely of an all-or-none flavor. Given the many ways in which discontinuous processes can mimic continuous ones, and vice versa, it is not evident that a deep theoretical resolution to this issue is forthcoming. However, from an educational perspective, it is worth remembering that the progression of learning may have as much to do with the nature of the task as with the nature of the learning process: Tasks in which information can be acquired incrementally are more likely to yield continuous learning functions than are tasks for which mastery hinges on the acquisition of a single principle or rule.

A related issue of continuity has been evaluated within developmental or lifespan approaches to understanding cognition: Does maturity induce qualitative changes in the mechanisms of learning? Indeed, much of the work in cognitive development and geriatric cognition can be characterized as a debate over which learning processes develop when, and which ones selectively decline with age. What is not in dispute is that there are qualitative changes in the patterns of results seen on tests of learning and memory. The elderly exhibit, for example, a much greater deficit relative to young learners on tests of recall than on tests of recognition. Fergus Craik has proposed that elderly learners will show, in general, deficient performance on tests that require them to self-initiate processing and relatively intact performance on tests that place little demand on such self-initiation. Tests of recall, which require the subject to develop and execute a systematic plan for addressing memory, retrieving relevant information, and outputting it successfully, play to the weaknesses of the elderly. Tests of recognition, which elicit only small deficits,

do not place a premium on self-directed processes but rather emphasize simple, presumably automatic, mechanisms.

Despite these results, it is not clear whether the underlying learning mechanisms actually differ between people of different ages, or if the apparent discontinuities arise epiphenomenally. Older learners may suffer more from proactive interference than younger learners, not because fundamental mechanisms differ, but rather because they have many more years of experience to interfere with learning and memory. If a constant amount of proactive interference affects different tests or tests for different types of information differentially, then empirical test dissociations can appear between age groups without a change in the underlying cognitive processes.

Aaron S. Benjamin

See also Learning; Learning Objectives; Learning Style

Further Readings

Craik, F. I. M. (1986). A functional account of age differences in memory. In F. Klix & H. Hagendorf (Eds.), *Human memory and cognitive capabilities* (pp. 409–422). Amsterdam: Elsevier.

Ebbinghaus, H. (1964). *Memory: A contribution to experimental psychology* (H. A. Ruger & C. E. Bussenius, Trans.). New York: Dover. (Original work published 1885)

Krechevsky, I. (1932). "Hypotheses" in rats. *Psychological Review*, *39*, 516–532.

Restle, F. (1965). Significance of all-or-none learning. *Psychological Bulletin*, *64*, 313–325.

COOPERATIVE LEARNING

Cooperative learning is an instructional process that engages students in collaborative discussions about the content to promote learning. The discussions may involve teaching, explaining, asking questions, quizzing, or checking, in an instructional activity where students actively share in the responsibility for learning. Cooperative learning processes significantly restructure classrooms from passive learning environments, with the teacher dominating the instructional conversation, into engaging environments where students actively participate in the learning environment.

Cooperative learning also attempts to change the social and motivational environment in the classroom to promote positive and supportive peer interactions and a positive orientation toward achievement and learning. This selection on cooperative learning will describe the philosophical and historical roots of cooperative learning. It will describe the theory behind the positive effects of cooperative learning and finally discuss some of the common cooperative learning methods used in elementary, secondary, and college instruction.

In many forms of cooperative learning, teachers initially lead instruction as a way to communicate new information or skills to students. As the students practice the new learning, the teacher guides them to develop more proficiency. Gradually, the students take the instructional lead as they interact with peers practicing collaboratively. This type of transfer of responsibility for learning, from the teacher increasing gradually to the students, is characteristic of most forms of cooperative learning.

There is an important distinction between cooperative learning and more traditional group work. Cooperative learning has structural features that are important to determining how the students work within the group and the effects that cooperative learning has on both academic and social outcomes. Most researchers believe it is important for well-structured cooperative learning to have a group goal and individual accountability. The group goal is the reason for the group members to collaborate; it motivates the students to work together and creates the interdependence necessary for a well-functioning group. Some examples of group goals include a written report, a product for a project, or an average test score for the group. The individual accountability is the reason for each group member to learn, and it is critical for the positive academic benefits found in cooperative learning research. The individual accountability ensures that each member of the group does his or her share of the work. Well-structured cooperative learning differs greatly from traditional group learning in large part because group work did not necessarily include individual accountability. For example, it is possible for one person in the group to write the whole report or to do most of the problems in the group activity. This kind of group work is less likely to lead to the kind of positive social and academic effects found in the research on cooperative learning.

Historical Background

Cooperative learning is not a new idea in education. Certainly, one of the early uses of cooperative learning occurred in the one-room schoolhouse, where one teacher was forced to teach students with a very wide range of abilities and ages. It is likely that teachers used collaboration among students as a pragmatic response to a challenging teaching situation. The philosophical notion of learning through peer collaboration is seen much earlier, in the writing of Quintilian (1st century) and Comenius (17th century), up to more recent work by John Dewey (20th century). All discuss the potential benefits of students teaching and learning from one another, yet it is unclear whether any of these earlier conceptions of cooperative learning took hold in the educational settings of the day.

Current applications of cooperative learning trace its development to sociology and social psychology in the mid-20th century, specifically to Gordon Allport's Social Contact Theory and Morton Deutsch's studies of group dynamics. While studying racial prejudice in social settings, Allport found that prejudice was reduced in settings where racially diverse people had close, substantive contact while working to achieve a common goal. The quality and depth of the racial interaction was an important factor in reducing racial prejudice. This became an important issue as public schools in the United States began the long task of desegregating, and overtly prejudiced behavior and poor peer relations were typical in newly desegregated schools.

Similarly, Deutsch's work provided social psychological support to this theory. In his work on competition versus cooperation, Deutsch found that in cooperative settings, where an individual's success was dependent on the success of others, individuals engaged in more positive communication with one another. These positive and supportive communication patterns led to groups with higher productivity and significantly more positive peer relations. This contrasted with the findings that competitive environments led to less group cohesion, fewer facilitative interactions, and generally less positive peer relations.

Social Outcomes of Cooperative Learning

These social psychological developments became of particular interest to educational psychologists in the 1960s and 1970s as research in desegregated schools found that racial prejudice and segregation within the schools were prevalent. Researchers applied social contact theory to the problem of diminishing prejudice and poor peer relations in newly desegregated schools. Early work on cooperative learning models such as Jigsaw and Teams Game Tournament attempted to put students together in groups to collaborate on common goals in an attempt to engage them in the kind of substantive contact that Allport had noted reduces prejudice. The research by Elliott Aronson, David DeVries, Robert Slavin, and David Johnson found that cooperative activities that engaged students of different races and backgrounds in substantive, academically oriented dialogue decreased prejudice and increased the quality of peer relations. The effects of cooperative learning on improving peer relations were found to transfer to relations outside the classroom, and positive peer relations remained during the school year even after the cooperative learning activities were over.

Academic Outcomes of Cooperative Learning

As cooperative learning research became more prevalent in schools, researchers noticed significant increases in academic performance among the cooperative groups. This was a natural extension of Deutsch's previous work on group dynamics as the group's positive and supportive communications led to higher productivity. Initially, the research on academic benefits used generic models of cooperative learning like Student Teams Achievement Division (STAD), Jigsaw, Learning Together, and Group Investigation (all described below). These models engage students in cooperative learning processes where they interact collaboratively on academic content. The models are not content specific and can be used with almost any instructional content. Typically, teachers use them as periodic activities to facilitate learning the content, often as an interactive way to practice the content or skills.

Similar research on student learning in university settings has also found academic benefits of cooperative learning. Researchers such as Alison King, Donald Dansereau, and Angela O'Donnell have found that peer collaboration during lecture and while reading textbooks can improve students' learning and retention of the content being presented.

Over time, some cooperative learning models became closely connected with specific content, becoming an instructional process integrated into daily instruction rather than an add-on activity that students engaged in periodically (e.g., weekly). Content-specific models of cooperative learning include Reciprocal Teaching and Cooperative Integrated Reading and Composition in reading and language arts, and Team Accelerated Instruction in mathematics. These content-specific models were found to have similar advantages for student learning.

The Nature of Cooperative Dialogue

For cooperative learning to be effective in promoting achievement, the cooperative dialogue must go beyond the "facilitative communication" first described by Deutsch. Research has found that not all student help is effective in increasing the achievement of both members of the dyad. Peer communications that result in terminal responses, where one student simply tells the answer to the other student, do little to increase the learning of either the student who gives the response or the student who receives the response. On the other hand, when a student provides an explanation as a response, like telling how to find the correct answer or explaining why an answer is correct, both students are likely to benefit from the experience. Receiving an explanatory response helps a student learn or encode something he or she previously did not know. More importantly, giving an explanation helps the other member of the dyad to process what he or she has learned in his or her own words, making more connections between the new information and his or her prior knowledge and enhancing understanding. Some research has suggested that the students who provide elaborative explanations actually experience greater achievement benefits than their peers who receive the elaborative explanations.

Multiple Theoretical Rationales for Cooperative Learning

A number of theoretical rationales have been used to develop and explain cooperative learning activities. As described above, the earliest rationale was based upon sociological theory relating to social contact and social psychological theory relating to group dynamics. These theories predicted and explained the positive social outcomes found in early cooperative learning research; however, the theories did little to explain the learning outcomes.

Generative Learning Theory

Perhaps the most prevalent theory for explaining cooperative learning's academic effects is generative learning theory. Generative learning suggests that when learners explain something to someone else in their own words, they increase their understanding of what they explain. Generative learning is situated within the information processing model of cognitive learning theory and specifically focuses on the processes involved in the activation of prior knowledge so that new knowledge is integrated with previously learned knowledge, thus increasing the probability of comprehension and recall of the new knowledge. This theoretical view explains the importance of giving elaborative explanations during cooperative learning to promote learning for not only the student who receives the explanation, but also for the student who gives the explanation. Generative theory provides a rationale for the evidence that high-ability students gain as much or more academically from cooperative learning as do average- or lower-ability students. Although common knowledge would cause one to expect lower-ability students to have the greatest benefits from cooperative learning processes, generative learning theory helps to explain why this is not necessarily the case.

The generative learning benefits during cooperative learning depend on students explaining or elaborating to one another; thus, teachers must monitor the interactions to make sure students provide explanations and do not provide terminal responses. Teachers also need to ensure that all students, regardless of ability, have an opportunity to provide elaborative explanations. To some extent, research has found that scripting interactions where students alternate roles in the elaborative dialogue can remedy the issue of equal opportunity to engage in the generation of explanations.

Sociocultural Learning Theory

Other cooperative learning research uses sociocultural theory and Lev Vygotsky's work to explain the academic effects of cooperation. Vygotsky suggests that development and learning occur as individuals internalize new information and skills, those within the proximal zone of development. In particular, Vygotskian theory states that for complex cognitive tasks, learners benefit

from interactions with more competent peers, like those interactions in cooperative learning. The theory suggests that interaction facilitates the internalization of newly learned skills. This type of interaction has also been called a cognitive apprenticeship, where learning occurs while engaging in academic interactions with a more competent peer or adult. The theory offers an understanding of the broader sociocultural context of cooperative learning, yet it may not fully explain the learning processes involved.

Sociocultural theory revolves around the ability of peers to provide guidance and feedback to one another during the collaborative dialogue, and their ability to do this effectively may depend on the age and sophistication of the students. Researchers have found that directly teaching, guiding, and monitoring students in how to engage in collaborative dialogue increases students' capability of providing these kinds of interactions.

Piagetian Learning Theory

Piagetian theory has also been applied to understanding the effects of cooperative learning, specifically through the concepts of construction of knowledge and cognitive conflict. The theory suggests that contradictory views, such as those that might occur in cooperation with a peer, create cognitive conflict. Piagetian theory suggests that cognitive conflict results in disequilibration that drives the learner to attempt to solve the internal conflict and hence construct meaning. Cooperative learning creates social interactions in which cognitive conflict occurs, and where continued collaborative dialogue (e.g., elaborative explanations) leads to conflict resolution and cognitive re-equilibration. The theory suggests that in cooperative learning, the interactions with peers stimulate this cognitive process that in turn increases learning of new information and skills.

This type of cooperative learning puts a premium on students learning how to work together in attempts to solve problems and resolve disagreements. Team-building activities are essential in order for groups to have sufficient cohesion and conflict resolution skills to work through their disagreements and collaborate in their investigations. The teacher's role is to promote team building and to act as a facilitator in guiding learning through investigation.

Sociocognitive Learning Theory

Albert Bandura's sociocognitive theory has also been used to explain the impact of cooperative learning processes on students' learning. Sociocognitive theory suggests that the learner will benefit from models in the environment, such as interacting with peers to promote both learning and motivation. At the same time, the theory suggests that the learner will reciprocally influence the group in the process, through goal setting, self-efficacy, and self-regulation during the cooperative processes. The reciprocal influence in this theory helps to explain many of the influences in peer collaboration; however, the internalization of learning, although mentioned, is not explicated.

To a great extent, these theories are not incompatible with one another, nor do they contradict one another in attempting to explain and describe the effects of cooperative learning on learning. Instead, they offer different foci in attempts to explain the nature of cooperative learning processes and the learning processes they promote. Because of their different foci, no single theory seems to capture all that is important about learning during cooperative interactions. Each offers a somewhat different insight into cooperative learning processes.

Selected Cooperative Learning Methods

There are a variety of methods for integrating cooperative learning into classroom instruction. The methods reflect differences in theoretical perspective, student population, or the nature of the instructional content. This list is by no means all inclusive; it is only intended to provide descriptions of methods to show the variety of different ways cooperative learning has been applied to classroom instruction. The initial models will be generic models that are cooperative learning processes that can be applied to a wide range of content. The latter models will be content-specific models specifically designed for use in one content area.

Jigsaw

Jigsaw is one of the earliest models of cooperative learning processes, and it was developed by Elliott Aronson. Jigsaw is best used with students in elementary school through college for learning narrative content (e.g., learning from a chapter in a text, doing a research report) and when the goal is content knowledge rather than skills. Teams are typically made up of four members. Members are assigned a portion of the

content that they are to research and learn so they can teach it to the other students in the team. In essence, they become the experts in that content. All of the students who are to become experts in the same content meet and work together on the important information for their area. After they have gathered the information, each expert returns to his or her team and teaches the content to the other students. In this way, each student has an opportunity to teach and elaborate on a portion of the content to his or her peers.

Student Team Achievement Division (STAD)

STAD is a cooperative learning method developed by Robert Slavin that is used in learning factual content (e.g., vocabulary, social studies or science information) as well as discrete skills (e.g., spelling, math computation, or language mechanics skills) for students in second through twelfth grade. Typically, it is used near the end of a unit of instruction and is used to promote active student practice in preparation for a test on the content. In STAD, the students are assigned to heterogeneous teams composed of four or five students. Initially, the teacher computes a base score in the content for each student, using a previous test(s) or a pretest. As students begin preparing for the end-of-the-unit test, they quiz one another about the material they are learning. Students then take the test, which is used to determine their grade and to determine their improvement points. Improvement points are based on the amount by which the student's performance increases above his or her previously computed base score. Improvement points are then used to calculate team scores for use in recognizing teams with good overall improvement. A large quantity of research on STAD consistently shows its positive effects on achievement and peer relations.

Teams Game Tournament

Like STAD, Teams Game Tournament (TGT) is used to promote students' learning of factual content or discrete skills, and is typically used near the end of a unit of instruction. In TGT, students from heterogeneous teams play an academic game, or tournament, that involves answering questions about the content, competing against three students of similar ability from other teams. The tournament involves students taking turns answering questions on the content. The other two students in the tournament can challenge the

student if they think he or she is incorrect. For every question the student answers correctly, the student gets a point. The points are used to calculate team scores, which, like STAD, are used for team recognition.

Group Investigation

Group Investigation is a cooperative learning method developed by Shlomo Sharan and Rachel Hertz-Lazarowitz in Israel that focuses on developing social skills and positive peer relations while learning academic content. The method is essentially cooperative inquiry, where students acquire, analyze, and synthesize information to solve a problem (e.g., write a research report, develop a plan of action). The groups work together to use resources such as texts, reference materials, and technology resources to collect relevant information. They then discuss how to organize the information for use in solving their specific problem. The teacher acts as a facilitator and guide, directing students to various information resources and asking them questions to guide their problem solving. At the end, each group presents a report of its work to the entire class. Research has shown that Group Investigation is particularly effective in increasing peer relations and developing students' interpersonal skills.

Learning Together

The Learning Together cooperative method developed by David Johnson and Roger Johnson is used in elementary and secondary school. The method emphasizes face-to-face interaction, positive interdependence, individual accountability, and interpersonal skills. Typically, Learning Together involves team building and teaching students appropriate interpersonal skills to facilitate the cooperative learning process. Some of the Learning Together methods have students study content together and quiz each other in preparation for individual tests. Other Learning Together methods involve students working together to complete a group test. Research on Learning Together has consistently indicated improved interpersonal relations and acceptance of peers.

Guided Reciprocal Peer Questioning

Guided Reciprocal Peer Questioning (GRPQ) is a cooperative learning method developed by Alison

King that was used initially with college-age students. Subsequent research has shown it can also be used effectively with children in the upper primary grades. GRPQ is used to help students actively process content presented in a narrative fashion, either in a textbook or from a lecture. Students are taught to ask questions about the content based upon question starters like "What does ____ mean?" "Describe ____ in your own words." "Explain why _____." Students ask a question to a partner, who attempts to answer it and then reciprocates by asking another question. This process can be used as students read a section of a textbook, or during a lecture when the instructor periodically stops to allow students to ask questions of a peer. Research has shown that students who use this method retain more information they read or hear in lecture, and it promotes metacognitive skills as students learn to ask themselves questions either during reading or when listening to a lecture.

Cooperative Integrated Reading and Composition

Cooperative Integrated Reading and Composition (CIRC) is a cooperative learning approach to teaching reading and language arts developed by Robert Slavin and Robert Stevens. CIRC has been used with students in second through fifth grade, and a companion model, Student Team Reading and Writing (STRW), has been used in middle school literacy instruction. CIRC is a multifaceted approach that involves students in learning both factual content (e.g., new vocabulary) and skills (e.g., reading comprehension and writing). The teacher provides initial instruction that is followed by students practicing collaboratively to complete tasks such as developing vocabulary knowledge, developing comprehension of the story, extending story comprehension through writing about the story, and engaging in the writing process for creative writing activities. Students check factual knowledge, make and elaborate on predictions, and provide clarifying explanations to one another about what they are reading. At the end of the instructional cycle (e.g., weekly), the students take tests, and the points earned on the test are used for team scores that are used for team recognition like that described above in STAD. Research studies have shown that CIRC and STRW have significant, long-term, positive effects on students' achievement, attitudes, and peer relations.

Reciprocal Teaching

Reciprocal Teaching is a cooperative learning method developed by Annamarie Palincsar to improve reading comprehension skills for students in elementary and middle school. Reciprocal Teaching begins with the teacher providing explicit instruction on comprehension strategies related to questioning, clarifying, summarizing, and predicting as students read. The teachers follow the initial instruction by guiding students as they practice using the strategies by prompting them with questions that they answer. Over time, the students take more responsibility by using questions to prompt one another; the questions are followed by answers in a reciprocal dialogue. The teachers monitor and guide the dialogue, helping the students increase their accuracy and proficiency in using the comprehension-fostering strategies. Research provides evidence of the efficacy of Reciprocal Teaching for students from third grade through middle school, with particular emphasis on struggling readers. Students also maintain the benefits of Reciprocal Teaching after the conclusion of its use in instruction.

Team Accelerated Instruction

Team Accelerated Instruction (TAI) is a cooperative learning approach to mathematics instruction developed by Robert Slavin for use in elementary and early middle grades. TAI focuses primarily on teaching math computation skills in a way that allows students to move at their own pace as they master each skill. The teacher begins the instructional cycle with an ad hoc instructional group based on those students beginning that particular unit. Following the initial instruction, the students engage in practice activities to develop mastery of the new skills. Students work in heterogeneous teams, allowing team members to give feedback and explanations to one another as they practice. There is ongoing progress monitoring as students take periodic tests to assess mastery. Students also earn points for their team by the number of instructional units they master and their level of performance on the mastery tests. Team points are used in determining team recognition like that in STAD, described above. Research indicates that TAI improves students' mathematics achievement and interpersonal relations for students of all abilities in second through sixth grades, and remedial secondary mathematics.

Across cooperative learning methods, there is remarkable consistency in how cooperative learning can positively influence academic and social outcomes in instruction across a variety of grade levels.

Robert J. Stevens

See also Cognitive View of Learning; Peer-Assisted Learning; Social Learning Theory; Vygotsky's Cultural-Historical Theory of Development

Further Readings

Johnson, D. W. (1999). *Learning together and alone: Cooperative, competitive, and individualistic learning.* Needham Heights, MA: Allyn & Bacon.

King, A. (1990). Enhancing peer interaction and learning in the classroom through reciprocal questioning. *American Educational Research Journal, 27*, 664–687.

O'Donnell, A. M., & King, A. (1999). *Cognitive perspectives on peer learning.* Mahwah, NJ: Lawrence Erlbaum.

Sharan, S. (1999). *Handbook of cooperative learning methods.* Westport, CT: Praeger.

Slavin, R. E. (1985). Cooperative learning: Applying contact theory in desegregating schools. *Journal of Social Issues, 41*(3), 43–62.

Slavin, R. E. (1995). *Cooperative learning: Theory, research, and practice.* Needham Heights, MA: Allyn & Bacon.

Stevens, R. J., & Slavin, R. E. (1995). The cooperative elementary school: Effects on students' achievement, attitudes, and social relations. *American Educational Research Journal, 32*, 321–351.

Van Meter, P., & Stevens, R. J. (2000). The role of theory in the study of peer collaboration. *Journal of Experimental Education, 69*(1), 113–127.

CORRELATION

Correlation is a descriptive statistical technique whereby the relationship between pairs of variables is assessed. The strength of the association between two variables can be determined either qualitatively, with a scatterplot, and/or quantitatively, with a correlation coefficient. A scatterplot is a two-dimensional graph with one variable plotted on each axis, whereby the slope of the least squares regression line (i.e., a linear "fit" through the maximum point cluster) indicates the overall direction and strength of the association. If the linear trend of ordered pairs is sloped upward (downward) from the origin and toward the right, the correlation between the two variables is positive (negative); no slope indicates a neutral relationship, and the variables are unrelated. The closer the data points cluster to this line, the stronger the association between the two variables.

The most common quantitative measure of correlation is the Pearson product-moment correlation coefficient, or simply Pearson's correlation (r), expressed as the ratio of the covariance between two variables to the product of their individual standard deviations. Pearson's correlation assumes a linear relationship for the variables under consideration and that the data are of ordinal or ratio scale and normally distributed. Correlation values range between -1.00 and 1.00, with more positive (negative) values indicating a stronger, direct (inverse) relationship and those values closer to zero signifying that the variables are not linked. The categorization of the strength of correlation coefficient values varies by discipline and inquiry. Typically, meaningful r values in the behavioral sciences are in the order of $|0.60 - 1.00|$ for strong, $|0.59 - 0.40|$ for moderate, and $|0.2 - 0.39|$ for weak associations. The researcher must evaluate the correlation coefficients with caution as an association between two variables does *not* equate with causation, and, in some cases, spurious relationships may be found between variables that defy logical explanation.

To test whether or not a correlation coefficient is significant, the t-test statistic is the most common measure used, especially for sample sizes smaller than 30. The t-test is the ratio of the Pearson's correlation coefficient (r) to the standard error of the estimate, a measurement of the variability of sample means. The t-test scores are often converted to probability values (p values), which identifies the probability of erroneously rejecting the null hypothesis if it is, in fact, true (i.e., a Type I error). Significant p values are typically those that are less than 0.05, indicating a 95% (or greater) confidence interval and, thus, the probability of making a Type I error less than 5%. In the case of a significant t-test, the null hypothesis that the two variables are independent (i.e., not correlated) must be rejected and the alternate hypothesis that the variables are related should be accepted. Based on the tail of the probability distribution, the alternate hypothesis is designated as either one-tailed (directional) or two-tailed (nondirectional), the latter being more appropriate if there is no a priori knowledge on the direction of the correlation between the two variables.

The use of Pearson's product-moment correlation coefficient in educational psychology is illustrated in the following example. Suppose a researcher wanted

to examine the relationship between the verbal test score portion on a standardized college entrance exam and number of hours spent in preparation for a group of students ($n = 50$). To interpret the correlation coefficient examining this relationship, the researcher should inspect both the sign (i.e., positive or negative) and the magnitude of the correlation. If the correlation coefficient is high and positive (e.g., .78), that means that those students who score high on the verbal section of the test are very likely to have spent more hours in preparation than other students. If the correlation is close to zero (e.g., .08), then there seems to be no direct relationship between the test score and how long the students prepared. Finally, if the correlation is high and negative (e.g., $-.62$), then the students who prepared more are not likely to do as well on the test as those who prepared less.

If the association between two variables is considered to be influenced by a third or more other variables, then partial correlation is employed. This technique is used to clarify bivariate relationships by removing (or partialling out) the confounding effects of other, possibly related variables. If the correlation coefficient between two variables remains unchanged while holding all other variables constant, then the control variables do not have any influence on the relationship. Conversely, large differences between the original Pearson's correlation and the partial correlation indicate that the relationship between the two variables is spurious and they are not directly linked, thus associated only by mediatory variables. In the previous example, the relationship between verbal test scores and study hours may be influenced by a third variable, such as whether or not English is the student's native language.

If the assumptions of Pearson's correlation are violated (i.e., not normally distributed, nonlinear relationship), a nonparametric correlation statistic, such as Spearman's rank [or Spearman's rho (ρ)] or Kendall's tau (τ) correlation coefficients, may be a more suitable alternative. Spearman's rho, the more common of the two methods, measures the strength of the association of ordinal (or reduced to ordinal) data by examining the ratio of the sum of the squared differences in the ranks of the paired data values to the number of variable pairs. Alternatively, Kendal's tau evaluates the linkage by examining the proportion of discordant pairs in a sample, where pairs are considered discordant if the product of the bivariate observations is negative. As with Pearson's correlation, coefficient values range

from -1.00 to 1.00 and are interpreted similarly, with larger values indicating a stronger relationship between the two variables.

In large samples, the coefficient values of both of these nonparametric test statistics are nearly identical; however, in the case of smaller sample sizes, Kendal's tau is usually the more efficient test.

Jill S. M. Coleman

See also Normal Curve; Statistical Significance

Further Readings

Gibbons, J. D. (1993). *Non-parametric measures of association.* Newbury Park, CA: Sage.
Goodwin, C. J. (2005). *Research in psychology: Methods and design* (4th ed.). New York: Wiley.

CREATIVITY

Creativity is among the most valued yet least understood of educational psychology constructs. For the purposes of this entry, *creativity* is defined as the interaction among aptitude, process, and environment by which an individual or group is able to produce *original* (unique, novel, unusual) and *adaptive* (useful, appropriate, meaningful) interpretations, ideas, behaviors, solutions, or products. Although most people view creativity and originality as synonymous, creativity scholars have emphasized the importance of including the additional criterion of *adaptiveness* in definitions of creativity. Without this added criterion, anything merely unusual would also pass as creative. Confounding creativity with uniqueness has, in part, resulted in the perpetuation of various misconceptions of creativity (e.g., creativity requires elimination of constraints and rejection of standards and conventions) and negative associations with those who are creative (e.g., creative individuals suffer from psychological disorders).

The Adaptiveness Criterion

Creativity researchers have demonstrated—in case studies of eminent creators and investigations of everyday forms of creative expression—the importance of including the criterion of adaptiveness in definitions of creativity. The adaptiveness criterion distinguishes creativity from novelty and provides a context for

evaluating particular contributions. Consider the every-day creativity of developing a new recipe. No matter how novel the combination of ingredients (e.g., oxtail, dandelion stem, maple syrup) or the technique (flambé)—if the resulting meal is inedible, then the recipe would hardly be considered creative. Rather, in order for the meal to be considered a form of creative expression, it would need to be original and useful.

The importance of adaptiveness is also illustrated in eminent forms of creative expression. Creativity scholars have documented that the highest forms of creativity typically require a deep understanding of the conventions and traditions of a particular domain: usually 10 or more years of formal study or apprenticeship. For example, eminent, creative jazz musicians (e.g., John Coltrane, Charlie Parker, Miles Davis) not only had an amazing ability to create highly original compositions, arrangements, and improvisations, but they were able to do so because of their deep knowledge and mastery of musical standards and instrumental techniques.

Sociocultural Context

In addition to recognizing the combined importance of originality and adaptiveness, creativity scholars also have stressed the importance of the sociocultural context in their definitions of creativity. This emphasis on the role of context has moved conceptions of creativity beyond traditional individualistic perspectives, which represent creativity as a static, genetically predetermined trait. Rather, most contemporary definitions of creativity emphasize the role that the sociocultural (and, at times, historical) context plays in determining what is considered original and adaptive. This is particularly the case with the most eminent creators. For example, case studies of great creators (e.g., Johann Sebastian Bach) highlight the role that the sociocultural context (e.g., the judgment of historians, critics, connoisseurs, and subsequent musicians) has played in determining the originality and adaptiveness of some creative body of work (e.g., Bach's musical compositions posthumously deemed highly creative).

The sociocultural context also plays a key role in more ubiquitous examples of creative expression. For example, in order for an architectural design to be considered creative, it must be judged to be not only original by the gatekeepers of the domain of architecture (i.e., professors of architecture, professional architectural societies and associations, historians,

fellow architects), but also adaptive (i.e., fit in with the standards and conventions of that domain, be structurally sound). Similarly, although a high school student's spoken-word poem may not be considered sufficiently original or adaptive to be included in the *Norton Anthology of Modern Poetry*, the student's poem might very well be considered creative within the context of her high school classroom, after-school poetry club, or local poetry event.

Common Misconceptions About Creativity

Like any construct that has wide appeal, popular beliefs about creativity have resulted in the development and persistence of several misconceptions. Three common misconceptions with particularly important implications for educators have resulted from confounding creativity with negative deviance, the arts, and eminence.

Creativity and Negative Deviance

The myth that creativity involves some degree of socially negative deviance is the upshot of believing that creativity represents a form of unconstrained novelty. In a classroom context, this misconception typically is represented in teachers' believing that if they encourage students' creative expression, they will ultimately take their class off topic by deviating from curricular standards, which will likely result in confusion or other classroom disruptions. It is not that teachers fail to see value in student creativity (in fact, most view creativity as important); rather, they struggle with how they can legitimately justify the incorporation of something in their standard curriculum that they believe will take instructional time away from learning the content. Not surprisingly, then, many teachers come to view creativity as extracurricular and not appropriate for their regular classroom instruction.

Creativity and the Arts

When most people think about creative expression, they often think of the visual or performing arts, and teachers are no exception. Consequently, teachers may struggle with the idea that creativity should be a common feature of their curriculum. If teachers believe that the visual or performing arts are the only form of creative expression, then it is easy to imagine

how creativity would come to be viewed as a rare, extracurricular luxury (e.g., having the students don costumes and perform the key elements from a historical description of the Salem Witch Trials, or having students write Haikus while studying the seasons). Although likely entertaining (and perhaps engaging), such representations of creativity have little to do with "creativity" as modern scholars conceptualize it.

Rather, many researchers highlight the importance of teachers helping students recognize and develop their creativity in various academic domains (e.g., scientific creativity, creative writing, artistic creativity). Incorporating creativity in the curriculum, from this broader perspective, would involve emphasizing the importance of developing novel ideas and products within the context and conventions of a particular academic domain. For example, a seventh-grade science teacher might support and encourage the development of scientific creativity in students' conceptualizations of their science fair projects. In order to do so, the teacher might require his or her students to develop a novel project idea (meeting the originality criterion of creativity) as well as adhere to the conventions and standards of scientific inquiry (meeting the appropriateness criterion of creativity). The creativity of such projects would then be judged within the sociocultural context of that particular science fair (by teachers, peers, guest judges, and other relevant stakeholders).

Creativity and Eminence

Finally, a common and damaging misconception is that only certain people can be creative. This belief is reinforced by the attention given to the revolutionary breakthroughs often attributed to a few creators (e.g., Wolfgang Amadeus Mozart, Mohandas Gandhi, Albert Einstein). It is true that the odds of a teacher having a classroom full of potential Martha Grahams or Martin Luther King, Juniors are slim to none. However, it does not then follow that teachers need not be concerned about supporting and enhancing students' creative potentials. As most researchers now recognize, all students have creative potential (although not necessarily on the level of magnitude of the rare, eminent creator). And although a few students will always demonstrate unusually high potential and creative ability and likely would benefit from additional curricular enrichment and acceleration, all students stand to benefit from having their creative potential recognized and supported.

Viewing creativity from a developmental perspective (with various degrees of creative magnitude) may help teachers recognize that all students have creative potential. For instance, creativity scholars have demonstrated how the creative life typically follows a trajectory starting with intrapersonal creative interpretations (referred to as mini-c creativity) moving to socially recognized forms of creative expression (little-c creativity), and, in very rare instances, going on to make revolutionary contributions that have lasting, historical, and cultural significance (Big-C Creativity). Big-C Creativity represents the highest forms of creative expression, whereas little-c represents the more ubiquitous, everyday forms of creativity. A great deal of scholarship has been focused on the Big-C level of creative magnitude and has yielded important and compelling insights. For instance, scholars who have conducted research on Big-C Creativity have demonstrated through case studies (i.e., in-depth studies of eminent creators) and histrometric methods (i.e., the quantitative, historical studies of general trends in creativity across time) that eminent levels of creativity result from a confluence of factors (e.g., motivation and skill, past and present environmental supports and constraints, the current and future sociocultural-historical context, and even luck).

The scholarship on Big-C Creativity provides an important backdrop for understanding the extremely rare and revolutionary impact that creativity can have on the development and advancement of science, technology, the arts, and societies as a whole. However, little-c creativity is often more relevant and important for educators to understand as it refers to the everyday forms of creative expression thought to be accessible to almost anyone (e.g., students designing a unique and accurate solution to an everyday science problem; students expressing their frustration about some life situation in poetry; students representing their understanding of some event of cultural significance in a unique and historically accurate narrative account, and so on). Indeed, nearly all students have a shot at producing what their teachers, peers, parents, and they themselves would consider to be novel and adaptive ideas, solutions, behaviors, and products. Research on this more everyday form of creativity has identified important cognitive strategies (e.g., divergent and evaluative thinking techniques) and environmental supports (e.g., creating engaging, intrinsically motivating environments) necessary for supporting students' creative potential.

The developmental perspective of creativity (which includes the various levels of creative magnitude) helps to dispel the myth that only certain people can be creative. At the same time, the developmental perspective acknowledges that although there will always be a few rare individuals (or groups) who have the capacity, environmental supports, and luck to attain creative eminence, the more everyday forms of creative expression are still important and attainable by almost anyone. For teachers, then, the goal becomes helping students to recognize and realize their creative potential within the various academic domains.

Theories of Creativity

Space limitations prevent a comprehensive review of creativity theories, but the trajectory of theoretical development can be characterized as moving from relatively focused, psychometric theories to multifaceted, relatively complex systems theories.

Psychometric Conceptions

The growth of creativity research in the second half of the 20th century was initiated by researchers studying creativity from a psychometric perspective. The work of J. P. Guilford and Paul Torrance was seminal in this regard. Guilford conceptualized creativity as a facet of his Structure of the Intellect model of human intelligence and developed a range of cognitive measures of creativity. Torrance, focusing more tightly on divergent thinking, developed the Torrance Tests of Creative Thinking, which remains one of the most popular creativity assessments and is used in several countries. Torrance's work popularized the conceptualization of divergent thinking in terms of fluency (total number of responses), flexibility (number of different kinds of responses), originality (the uniqueness and rarity of responses), and elaboration (the detail of responses).

Systems Theories

Starting in the 1980s—and mirroring theoretical developments in other fields—creativity researchers began to question the utility of psychometric conceptions of creativity. These scholars noted that creativity was a broad construct with multiple, diverse influences and outcomes (as opposed to the more narrow cognitive conceptions of the traditional psychometric

approach). This changing attitude resulted in part from the broadening of the field to include researchers with wide-ranging interests and backgrounds. The result was a flurry of systems theories of creativity from the late 1980s to the mid-1990s. Mihaly Csikszentmihalyi's systems theory, for example, emphasizes the interaction between an individual, the domain, and the field in which creativity occurs. The systems perspective, common in most systems theories, focuses on the confluence of events that occurs when creativity happens rather than considering creativity to be a trait of an individual. For example, most people do not understand Einstein's Theory of Relativity, but accept it as true because the gatekeepers in the field accept it. A domain is the cultural or symbolic portion relevant to creativity. Creativity "happens" when an individual changes a domain. Other popular and influential systems theories include Teresa Amabile's Social Psychology of Creativity, Robert Sternberg and Todd Lubart's Investment Theory of Creativity, and Mark Runco's Psychoeconomic Theory of Creativity, among others.

Types of Creativity Research

Scholars representing multiple disciplines (e.g., psychology, education, history, the arts, sociology, anthropology, and business) have contributed to the development of creativity theory, research, and practice. The upshot of this multidisciplinary effort has resulted in the broadening of the field and the emergence of multiple avenues and venues for creativity research. For instance, the broadening of the field is evinced in a steady stream of new creativity handbooks and edited volumes, the sustainability of multidisciplinary creativity research journals (e.g., *Creativity Research Journal, Journal of Creative Behavior*), and the emergence of new scholarly journals (e.g., *Journal of Thinking Skills and Creativity; Psychology of Aesthetics, Creativity, and the Arts*).

Several distinct approaches are used to examine creative phenomena, with a majority of research historically relying on *psychometric methods*—the direct measurement of creativity and its perceived correlates in individuals. Practically all current work on creativity is based upon methodologies that are psychometric or were developed in response to perceived weaknesses of creativity measurement. Work by Guilford and Torrance, which, as mentioned earlier, focused on cognitive assessment, formed the foundation for

psychometric approaches, and much of their work remains in wide use among creativity researchers. But psychometric methods have been extended to the study of attitudes, personality, motivation, and other noncognitive aspects of creativity. The advantages of psychometric approaches include the existence of a number of refined instruments, which promotes comparison of results across studies; disadvantages include the often narrow lens through which creativity is viewed and the lack of psychometric advances to match the considerable theoretical advances in recent decades.

The *experimental approach* is quite similar to the psychometric approach in that experimental researchers use many of the same instruments used by psychometricians to measure creativity. Perhaps the best known line of research in this area involves Amabile's studies of the impact of reward and evaluation conditions on creative production. Although experimental designs include added logistical and financial costs (as compared to nonexperimental designs), the advantages of the experimental approach include the confidence with which conclusions can be made about intervention effectiveness. Still, as some scholars have noted, because experimental designs require direct intervention and manipulation (as opposed to naturalistic observations and case studies), they may preclude the ability for documenting how creativity develops (more naturally) in a given sociocultural context.

The *historiometric approach* involves the analysis of quantitative data drawn almost exclusively from historical documentation. Although studies of eminence have relied on historiometric techniques at least as early as the 1950s, this empirical approach is best exemplified by the prolific works of Dean Keith Simonton, who has both created and refined the methodologies and used these techniques to study creativity exhaustively. Historiometric methods have been used to study the relationships between creativity and leadership, age, productivity, invention, and eminence, among many other areas. An advantage of this approach is the ability to consider creativity within its proper historical context, which proves to be difficult using other empirical approaches. On the downside, this approach relies almost exclusively on the study of Big-C Creativity, potentially limiting its application to everyday creativity.

The *biographical* or *case study approach* is similar to the historiometric approach in its focus on eminent creators and their work. However, the case study approach strongly favors qualitative research methodologies, with the attendant advantages regarding depth of analysis and limitations regarding generalizability. For both historiometricians and biographers, their focus on Big-C Creativity (i.e., indisputable instances of creativity) helps researchers avoid the problems and issues surrounding the definition of creativity faced by investigators adhering to most other methodological approaches. But again, applicability of results to everyday creativity may be an issue.

Although the use of *biometric methods* to study creativity, such as functional magnetic resonance imaging (fMRI), has significant potential, their application to the study of creativity has been highly limited. The ability to study brain function during creativity problem solving has clear advantages, but the use of fMRI in this context has been limited because of the high cost of such research, the infancy of the technology, and a lack of interest among neuropsychologists. As these issues change—and evidence suggests that this is happening—biometric approaches should become more common.

Although the various approaches to the study of creativity share many features, fundamental differences distinguish them from one another. The most significant differences involve the research designs commonly used, specific areas focused upon by researchers, and the time frame in which data are collected.

Major Empirical Themes

Two lines of inquiry are important for educators and educational psychologists to understand: the issue of domain specificity versus domain generality, and issues surrounding the enhancement of creativity.

Generality Versus Specificity

Whether creativity is domain general or domain specific is one of the hottest—yet most enduring—debates among creativity researchers. In essence, the debate boils down to whether people who are creatively productive in one domain are likely to be similarly productive in a different domain (or domains). Such distinctions are important because the generality position implies that general skills can be used to enhance creativity. But if creativity is domain specific, general creativity training would be highly

unlikely to promote transfer across domains and would, in effect, be pointless.

The specificity position, which is currently the dominant perspective, is supported by recent research on situated cognition and other areas within cognitive science. For instance, domain-specific creative tasks tend to have low correlations with each other, and task-specific training typically yields task-specific creative performance. However, other researchers argue that a domain-specific conception fails to account for theoretical and methodological issues. For example, a systems approach to creativity takes the view that creativity cannot be determined entirely within a domain. Implicit theories of creativity (i.e., layperson or folk definitions) also support a domain-general perspective.

A more fruitful question focuses on which aspects of creativity are domain or content general and which are domain or content specific, with the answer changing as a person develops and matures. As people age, they are forced to make choices, related to both their careers and their personal lives, that necessarily limit their ability to work productively in multiple, distinct domains. Some creative individuals do work in multiple domains, but they tend to be the rare exceptions that prove the rule. The implications of this perspective for practice are that some aspects are domain general, some are domain specific, and enhancement efforts should focus on helping individuals develop their creative abilities in whichever areas they choose—keeping in mind that some of the skills and attitudes will transfer to other domains and tasks.

Enhancement

Research on the effectiveness of efforts to enhance creativity is controversial, conflicting, and often disappointing. For example, many studies provide evidence of the ability of divergent thinking techniques to increase divergent thinking test scores, but these gains are often short-lived, and the skills prove difficult to apply to other contexts. Another example involves group problem solving. The work of Alex Osborn, Torrance, and many others suggests that groups generate more ideas than individuals working under otherwise similar conditions, and that group participants rate the creative performance more favorably than individuals working on similar problems. At the same time, many researchers have found that group problem-solving strategies, such as brainstorming, produce fewer and less original ideas than do individuals. The creativity enhancement literature is rife with such contradictions.

Several researchers have suggested possible reasons for the lack of clear effectiveness information regarding creativity enhancement: the lack of appreciably new enhancement strategies at the same time that our theoretical understanding of creativity has advanced exponentially; widely held misconceptions and myths about creativity (often created by an over-reliance on Big-C Creativity), which most "classic" enhancement efforts fail to address; negative effects of the numerous pop psychology and commercialized approaches (many of which are merely repackaged versions of early divergent thinking strategies); early work conducted in relative isolation from mainstream psychology; reliance on strategies and programs that assume a strong domain-general position; and conflicting or vaguely implied definitions of creativity. This last point is very important—if researchers studying creativity enhancement fail to define creativity or define it imprecisely, conflicting results should be expected.

Looking at creativity enhancement through the lens of the definition of creativity described earlier, a promising avenue for enhancing creativity is to optimize the time an individual spends in ideal aptitude-process-environment contexts. This may be similar to Csikszentmihalyi's concept of flow and Sternberg's emphasis on contextual awareness in his theories of intelligence. However, the ideal creative context needs to be identified for each individual. Not only will this vary across individuals, but across time, process, and environment as well. Many approaches to problem-based learning promote these principles, suggesting that related activities may be a fruitful avenue for future enhancement efforts.

Joseph Renzulli's Schoolwide Enrichment Model emphasizes many of these approaches and may be the most appropriate program for the enhancement of creativity. Initially designed as an education system for gifted students, Renzulli's approach has been described by some observers as a creativity enhancement model, and viewed from this perspective, it is considered to be quite successful. Renzulli has recently adapted the model to fit a variety of instructional contexts, and his work should be consulted by anyone considering large-scale enhancement efforts.

In the end, it is reasonable to conclude that a growing number of educational psychologists value creativity, would like to enhance it, and are confused

about how to proceed. Addressing complex phenomena such as creativity is never easy, and recent conceptual and empirical advances should illuminate the issues surrounding enhancement of creativity.

Creativity and Educational Psychology

Given that all students have creative potential and that creativity can help complement and extend learning within the various academic domains, there should be little doubt that creativity represents an important construct for educational psychologists. Although creativity has dwelled somewhat on the margins of educational psychology (particularly during the era in which behaviorism held sway over the field), the rise in constructivist views of learning has established a growing recognition that creativity and knowledge construction are complementary processes. This recognition is not entirely new, as many early progenitors of modern constructivist views of learning (including John Dewey, Jean Piaget, and Lev Vygotsky) made explicit references to the relationship between learning, creativity, imagination, and the development of robust understanding. What is new, however, is the wealth of knowledge both in the domain of educational psychology and in the interdisciplinary field of creativity studies.

For instance, educational psychologists have established important theories and empirically tested insights pertaining to student motivation, learning and development, assessment, the psychology of teaching, and contextual factors that mediate effective teaching and learning. Moreover, advances in research methods (both qualitative and quantitative) and advances in analytic techniques (e.g., everything from hierarchical linear modeling to more situated, interpretive analyses) position educational psychologists to make significant contributions to the study of creativity and its relationship to teaching and learning. Similarly, creativity researchers have engaged in parallel efforts in which they have examined the development and expression of creativity in educational contexts. For instance, creativity researchers have developed full programs of research (including the development of theoretical models and innovative analytic and assessment techniques) aimed at examining the motivational, cognitive, affective, and socio-environmental factors that mediate the fulfillment of students' creative potential. What is now needed is a well-articulated bridge between these parallel efforts, both to reduce unnecessary redundancies and to maximize the synergies between these related knowledge bases and lines of inquiry. Educational psychologists have a unique opportunity to bridge these complementary programs of research such that existing knowledge can be brought to bear on the advancement of what is known regarding the relationship between human learning and creative expression.

Jonathan A. Plucker and Ronald A. Beghetto

See also Cognitive View of Learning; Constructivism; Gifted and Talented Students; Motivation

Further Readings

Beghetto, R. A., & Kaufman, J. C. (in press). Toward a broader conception of creativity: A case for mini-c creativity. *Psychology of Aesthetics, Creativity, and the Arts.*

Houtz, J. C. (Ed.). (2003). *The educational psychology of creativity.* Cresskill, NJ: Hampton Press.

Kaufman, J. C., & Bear, J. (Eds.). (2006). *Creativity and reason in cognitive development.* New York: Cambridge University Press.

Plucker, J. A., Beghetto, R. A., & Dow, G. T. (2004). Why isn't creativity more important to educational psychologists? Potentials, pitfalls, and future directions in creativity research. *Educational Psychologist*, *39*, 83–96.

Runco, M. A. (Ed.). (1997). *Creativity research handbook, Vols. 1, 2, and 3.* Cresskill, NJ: Hampton Press.

Runco, M. A., & Pritzker, S. R. (Eds.). (1999). *Encyclopedia of creativity.* San Diego, CA: Academic Press.

Sawyer, R. K. (2006). *Explaining creativity: The science of human innovation.* New York: Oxford University Press.

Sternberg, R. J. (Ed.). (1999). *Handbook of creativity.* New York: Cambridge University Press.

Sternberg, R. J., Grigorenko, E. L., & Singer, J. L. (Eds.). (2004). *Creativity: From potential to realization.* Washington, DC: American Psychological Association.

CRITERION-REFERENCED TESTING

Definition and Features

A criterion-referenced test (CRT) provides a measure of an individual's absolute performance or behavior on a well-defined domain. The domain may include

a set of learning/behavioral objectives to be mastered or a set of standards to be achieved. The history of CRTs dates back to a 1963 essay by Robert Glaser in which he introduced criterion referencing as a new type of approach to test development and interpretation. Glaser indicated that the absolute comparisons that formed the basis of CRTs were preferable to the relative comparisons made by the norm-referenced tests (NRTs) widely used at that time.

One way to gain a better understanding of exactly what a CRT entails is to compare it to an NRT. Within this section, CRTs are further defined and comparisons to NRTs are made. The next section discusses the development, use, and interpretation of CRTs in large-scale tests used for accountability purposes. It includes issues of alignment and validity, performance levels, and standard settings. The third section discusses the use of CRTs for classroom instructional purposes. The final section briefly describes guidelines available to assist test users in selecting and interpreting CRTs.

What Does a CRT Measure?

Whereas NRTs compare an individual's performance to the performance of others in a comparison group, such as all students in a state or across the nation, CRTs compare an individual's performance to standards or learning objectives. Because of this distinction, interpretations of NRTs are often referred to as relative comparisons and CRTs as absolute comparisons. In a CRT, test users are concerned about knowing whether the student has achieved a standard or mastered a learning objective. They are not concerned about the position or ranking of the student relative to other students.

What Is the Domain for a CRT?

CRTs consist of a well-defined domain of knowledge, skills, and/or behaviors to be measured. The domain for a CRT is narrower and more clearly delineated than the domain for an NRT, which is usually broader and covers many objectives. The level of specificity, however, varies somewhat across different types of CRTs. In some classroom situations, CRTs are called *objective-referenced* because they measure detailed learning outcomes (e.g., adding two three-digit numbers that require carrying). In large-scale tests that are standards-based, the criteria may be slightly less detailed (e.g., measuring whether a student can represent mathematical situations using algebraic symbols).

How Are the Items Designed?

In NRTs, items are usually developed so that the difficulty level is average and discrimination among student scores is high. Because easy items do not lend themselves to discriminating among individuals, they are not usually used on NRTs. However, for CRTs, the item difficulty or discrimination is not of utmost importance. Rather, the most critical aspect of developing a CRT is that each item must have a direct match to a learning objective, behavior, or standard within the domain. Depending on the purpose of administering the CRT, the items may be very easy or difficult. When a test is given immediately after instruction to determine whether students attained the knowledge necessary to move on to the next topic, it is likely that the items will have a low difficulty level. Also, there would be no concern about whether the test was able to discriminate among students.

What Types of Scores Do CRTs Produce?

Unlike NRTs, which provide standard scores and percentiles to describe a student's placement or ranking within a comparison group, CRTs provide percentages that indicate whether a student has achieved mastery or proficiency regardless of the performance of other students. CRTs may indicate (a) the number or percentage of items that were answered correctly by each individual, (b) the speed of performing a certain task (e.g., typing on a keyboard), (c) the precision of performance (e.g., measuring an object to the nearest millimeter), or (d) the quality of a behavior (e.g., tying a shoe). Interpretations from these scores help to determine whether a student has achieved mastery on a set of skills, tasks, or behaviors. Other CRTs, such as those that are standards-based, combine a student's performance on a set of items and categorize him or her into one of several performance or achievement levels. The student might be classified as being basic, proficient, or advanced with regard to a standard in a subject area. Results may also be aggregated across all students in a school, grade, or subgroup (e.g., ethnicity, gender) so that teachers and administrators can examine the percentage of students classified within each level.

What Are the Purposes for Using CRTs?

CRTs can be used for several purposes. It is the responsibility of the test user to determine whether the test design, item design, and type of scores match the intended use and interpretation of the test. One set of purposes for using a CRT is to determine whether individual students have mastered the learning objectives, acquired the skills necessary for placement into a particular program, or demonstrated the ability to perform essential behavioral tasks. In these situations, reporting meaningful score interpretations to parents is of extreme importance because it directly relates to their child's success. Another purpose of a CRT is to inform classroom instruction and ensure that adequate learning experiences are being provided to students. Teachers can identify students' strengths and weaknesses and modify instruction to address these areas.

CRTs may also be used in program evaluation. They provide answers to questions such as, "How effective is a new instructional program in our school?" "What impact does the program have on our students?" and "Is it worthwhile to continue the program?" This information is valuable to teachers and administrators as they judge the effectiveness of a program or curriculum in their school. Finally, another use of CRTs in recent years is to hold schools accountable for student learning. Within this era of No Child Left Behind (NCLB), state testing programs are using large-scale CRTs to determine whether students are achieving standards for reading, mathematics, and science. Although the media often rank schools in terms of student performance, it is not helpful for a school to know only that they are performing better or worse than other schools. For accountability purposes, the school needs to know the percentage of students achieving the state standards so they can work toward having every student perform at least at the proficient level. The following section contains more information on large-scale CRTs.

Large-Scale Criterion-Referenced Testing

In the 1990s, there was a movement by states to develop academic content standards and create state-level criterion-referenced assessments to measure student performance against those standards. Many assessments included performance-based tasks, and a few states, such as Maryland, had tests that were entirely performance based. When the No Child Left Behind legislation was enacted, it became necessary for all states to have an assessment system that would identify whether students were proficient. One of the major goals of NCLB is that all students will be proficient in reading and mathematics within a period of 12 years.

Because of NCLB, some states had to change their assessment systems, especially those with tests that were completely performance based. Maryland's test, for instance, provided school and grade-level results but not individual student scores, which are necessary for NCLB. In order to provide student scores, the number of items on which each student tests must be large enough to ensure a stable estimate of his or her performance. Therefore, Maryland had to abandon its performance-based assessment and instead use one that incorporated additional items that were not performance based. Many other states created new assessments; modified their existing assessment; or reverted back to using an off-the-shelf, commercially published criterion-referenced test. The latter is somewhat undesirable because the test may not provide an exact match to the state standards. However, test publishers are now beginning to collaborate with state representatives to produce custom-designed tests that will meet their specific needs.

Alignment and Validity

One of the major features of a well-designed CRT is that the test items are directly matched to a specific domain, standard, or learning objective. Large-scale CRTs used for accountability purposes are considered to have high stakes. For this reason, it is imperative that there exists a high degree of alignment among all aspects of the assessment system—state standards, test items, curriculum taught in school district, and instruction in teachers' classrooms. If any of these aspects is not in alignment, then the validity of the interpretations made from the CRT results is suspect.

Several models are available to examine alignment between state standards and tests. The models vary in their complexity. Low-level models involve experts examining the content of each test item and determining the degree to which it is related to the content in the standard. Models at the next level have additional criteria for determining the match between items and standards, such as the cognitive complexity of the item. The most complex alignment models examine several interrelated dimensions, including content, depth, emphasis, performance, and accessibility.

Even when there is a high degree of alignment between standards and test items, if there is not a match between the test and classroom instruction, then the assessment system will not be valid. Students must be given an opportunity to learn the content and processes specified in the standards. If they do not have this exposure in the classroom, they cannot be expected to demonstrate what they know and can do. Moreover, the conditions under which they solve problems should be the same in the testing situation as in the classroom; for instance, the accessibility of manipulatives and calculators in math. Various approaches have been described by educational researchers to examine the match between the test and classroom instruction. Some approaches collect evidence from questionnaires and interviews with teachers and principals. Teachers rate the similarity between the content and processes they teach in the classroom to the specific state standards. These self-report data also gather information on the unintended or potentially negative consequences of the assessment system, which is another important type of validity to be examined for CRTs. A more direct technique for gathering evidence of alignment is through classroom observations or the collection of sample instruction and assessment materials used by teachers. Classroom materials can be evaluated using a coding scheme that focuses on the content, level of cognitive complexity, and item response types, and comparing these to the test items.

Performance Levels

One way that standards-based CRTs report scores is in terms of a small number of performance levels. The number of levels is typically between three and five. The Pennsylvania System of School Assessment (PSSA) uses four performance levels: below basic (inadequate academic performance), basic (marginal), proficient (satisfactory), and advanced (superior). For the purposes of NCLB, each student is classified into one of the levels for each subject area assessed—reading and mathematics. For the purposes of disseminating test results to other stakeholders, Pennsylvania produces individual student reports for parents and school summary reports for schools and districts.

An individual student report for PSSA math identifies the performance level that the student achieved when aggregating results across items on the math assessment. Most items are multiple choice and assess a variety of skills from recall to problem solving.

A small set of open-ended items is included to measure students' problem-solving and reasoning skills and to require them to explain answers, describe solution strategies, create graphs, and so on. The individual student report also provides results for each academic standard. For the 2005 PSSA, there were five major reporting categories in math: numbers and operations, measurement, geometry, algebraic concepts, and data analysis. They are aligned to the content standards used by the National Council of Teachers of Mathematics. For each category, the number of points achieved by the student is shown in comparison to the total number of points possible. As an example, a student might receive 12 out of 14 possible points for the geometry standard that assesses knowledge of shapes, properties of shapes, and use of geometric principles to solve problems. This detailed information helps parents identify their child's strengths or weaknesses in math.

A school summary report for mathematics shows the number of students taking the assessment and the percentage of students in every grade level scoring at each of the four performance levels. With respect to individual standards on the assessment, the report lists the number and percentage of total points achieved by all students in a grade level. This allows teachers and school administrators to interpret results for the purposes of examining the math curriculum and planning instructional improvements.

It should be noted that even though the PSSA is a criterion-referenced test, the individual report and school summary reports also provide norm-referenced interpretations of results. The individual report presents average state-level scaled scores, gives a percentile to show the percentage of students across the state that performed at the student's scaled score or lower, and charts the difference between the student's points achieved on each standard compared to the state average points achieved. Similar norm-referenced interpretations are made in the school summary report. As Robert Linn and Norman Gronlund indicated in their book, *Measurement and Assessment in Teaching*, the distinction between large-scale CRTs and NRTs is becoming less clear-cut. A dual interpretation of tests is now on the increase.

Standard Setting

A natural question that follows a discussion of performance levels is about the criteria used to place

a student into a level. Usually, cut scores between levels are established. The process of creating the cut scores is called standard setting. There are many different methods of standard setting, all of which require making judgments. It is for this reason that some people have questioned the validity of performance levels, saying that it is difficult to get judges to agree on the cut scores, and thus they are inherently arbitrary. However, others believe that consistency among judges can be obtained through rigorous training and the use of previous test score data. Recent research is focusing on comparisons of standard-setting methods and their validity.

The Bookmark standard-setting method is widely used in kindergarten through twelfth-grade settings. The PSSA cut scores are based on a modification of this method. Judges, who were educators across the state, were given detailed written descriptions of each performance level and a booklet of all items on the assessment ordered from the easiest to most difficult based on previous students' responses. The judges were asked to think about a student who would be at the borderline between two levels (e.g., below basic and basic) and assess whether the student had a high probability of answering the item correctly. Each judge recorded his or her decision. Discussions occurred after each round, and the process was conducted repeatedly. After the judges' final recommendations were received, student scores were converted to scaled scores that identified a range for each level. For example, in 2005, a fifth-grade student with a scaled mathematics score between 1312 and 1482 would be categorized as proficient in math. A fifth-grade student with a scaled math score equal to or above 1483 would be advanced.

CRTs Used in Classroom Instruction

In Anthony Nitko's 2004 book, *Educational Assessment of Students*, he describes three types of referencing frameworks used to interpret achievement: norm-referencing, criterion-referencing, and self-referencing. Each of them contains a different perspective on achievement while increasing the meaningfulness of test results. A norm-referencing framework is not usually the most useful in making instructional decisions in the classroom or informing parents. When parents know only that their child is better at math than other students, they are left wondering about the kinds of math problems their child can solve successfully.

When you want to identify specific learning objectives that are achieved by each child, a criterion-referenced framework is most appropriate. Teachers can uncover math processes that are difficult for students (e.g., transforming numerical data to a visual display), and they can identify specific misconceptions students may have (e.g., about interpreting a line graph showing the relationship between speed and time). Teachers can learn about students' thinking and reasoning on complex math problems, and they can gauge their students' ability to communicate mathematically. All of this knowledge that the teacher learns about students becomes useful in planning and modifying instruction for all students. The teacher may decide to reteach, review, or proceed ahead with a new topic. Knowledge gained about individual students' deficiencies can also aid the teacher in individualizing instruction, giving each student additional learning opportunities or resources that address his or her specific needs.

A criterion-referencing framework provides summative and formative feedback. The previous paragraph gave examples of evaluating students for formative purposes in order to continually improve the teaching and learning process. The framework also allows for summative evaluations that inform teachers and parents about the level of achievement at which students are performing in relation to the absolute learning objectives. It is possible that all or most of the students in a class achieved mastery or are classified at the highest performance level. In this case, the items on the CRT would be considered "easy" for students, which is acceptable if the test was intended to measure knowledge necessary for proceeding with further instruction. It is also possible that all or most students do not achieve mastery or are not classified as proficient. Several reasons might explain this outcome. The learning objectives in the classroom may not be aligned with the items on the assessment, or the objectives may not be placed appropriately in the curriculum. Another reason may be that students did not have an opportunity to learn the skills and knowledge because the instructional methods and techniques used by the teacher were not effective. Teachers can reflect on these possibilities and modify accordingly.

Before leaving this section, a brief mention is made about the third type of referencing framework identified by Nitko because it serves to illustrate what a criterion-referencing framework is *not*. A self-referencing framework is based on the teacher's perceptions of a student's capacity to perform well in

a content area. Students would receive a high grade if they are achieving at or above the level at which they are perceived to be capable. Likewise, the reverse would be true, that is, students would receive low grades if they are achieving below the level at which they are capable. A student may also receive a high grade if he or she began the learning process with little or no prior knowledge and then made some improvement. As would be expected, there are pros and cons to this framework. Proponents believe that it can help increase student motivation and reduce competition among students. Those not in favor say that teachers' perceptions of students' capabilities are subjective and also that the framework leads to a grading procedure based on effort alone. Furthermore, it may produce a ceiling effect for those students who come to class with a great deal of prior knowledge.

Criterion-Referenced Letter Grades

There are several ways to assign a grade based on a variety of evaluation components. Nitko describes several methods, three of which are summarized here. One method involves using a fixed percentage. First, the number of points earned on each evaluation component used in the classroom is converted to a percentage. The percentages are then transformed to letter grades. The teacher defines a range of percentages for each letter grade. The ranges are equivalent across components.

The total points method first produces the total number of points a student achieved on all of the evaluation components. The number of points achieved is compared to the maximum number of points available across all components. Ranges of maximum points are defined for each letter grade. In this method, the number of points assigned to each component determines the weight, or worth, of each component. For example, a project might be assigned 50 points, whereas a short quiz might be assigned 10 points. The third method is called the quality level method. It is similar to using a rubric for scoring student responses on a performance task. The quality of student performance necessary to achieve each level is thoroughly described. Teachers match the quality of student work to the appropriate letter grade.

Guidelines for Appropriate Selection and Use of CRTs

This final section describes two external sources that serve as guides for educators and psychologists who develop, select, and use assessments. One source is called the *Standards for Educational and Psychological Testing.* The most recent version was developed in 1999 through a joint effort by the American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education. The document identifies standards that are designed to evaluate educational and psychological tests. These widely accepted standards and criteria provide information to test developers, users, and takers on the entire testing process. The document is organized into three parts: (1) test construction, evaluation, and documentation; (2) fairness in testing; and (3) testing applications.

Another source is the *Code of Fair Testing Practices*, published by the national Joint Committee on Testing Practices in 2004. The Code applies to all educational assessments regardless of their referencing framework (CRTs or NRTs) or the mode in which they are administered (computer-based, paper-pencil, or performance). The overall purpose is to ensure that tests are fair to all test takers regardless of their demographics (e.g., ethnicity, gender, age) in terms of a variety of aspects, including the standardization of administration conditions, preparation for the test, knowledge about the test's content and purpose, accurate reporting of results, and appropriate interpretation of the results.

The Code was designed to represent selected portions of the *Standards for Educational and Psychological Testing* by making the information relevant and meaningful to states, districts, schools, organizations, and individuals. Even though the Code is specifically written for professionally developed tests, the guidelines can also be used by teachers to inform and improve the fairness of their classroom assessments. The Code is organized according to four areas: (1) developing and selecting appropriate tests, (2) administering and scoring tests, (3) reporting and interpreting results, and (4) informing test takers. In each area, guidelines are given for both test developers and test users. Examples of guidelines in the third area are that test users should always take into account the nature of the content assessed when making meaning of test results, ensure that the test is used for its intended purposes, gain knowledge of the type of scores reported (e.g., method for establishing cut scores for mastery or performance levels), and understand that one test score should not be used as the sole determinant in making a decision about a student.

Carol S. Parke

See also Grading; High-Stakes Testing; Learning Objectives; No Child Left Behind; Norm-Referenced Tests

Further Readings

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing.* Washington, DC: Authors.

Bhola, D. S., Impara, J. C., & Buckendahl, C. W. (2003). Aligning tests with states' content standards: Methods and issues. *Educational Measurement: Issues and Practice*, *22*(3), 21–29.

Data Recognition Corporation. (2005). *Technical report for the Pennsylvania System of School Assessment: 2005 reading and mathematics.* Maple Grove, MN: Author.

Glaser, R. (1994). Criterion-referenced tests: Part I origins. *Educational Measurement: Issues and Practice*, *13*(4), 9–11.

Joint Committee on Testing Practices. (2004). *Code of fair testing practices in education.* Washington, DC: Author.

Karantonis, A., & Sireci, S. G. (2006). The Bookmark standard-setting method: A literature review. *Educational Measurement: Issues and Practice*, *25*(1), 4–12.

Linn, R. L., & Gronlund, N. E. (2000). *Measurement and assessment in teaching* (8th ed.). Upper Saddle River, NJ: Merrill/Prentice Hall.

Nitko, A. J. (2004). *Educational assessment of students* (4th ed.). Upper Saddle River, NJ: Pearson.

Osterlind, S. J. (1988). *Using CRTs in program evaluation.* National Council on Measurement in Education Instructional Module. Retrieved February 4, 2007, from http://www.ncme.org/pubs/items/9.pdf

CROSS-SECTIONAL RESEARCH

Research is commonly designed to test causal hypotheses. Whether the hypothesis is general or specific, it can be expressed as "IV causes DV," where IV is an independent (causal) variable and DV is the dependent (consequent) variable. There are three generally accepted conditions for establishing causality— *temporal ordering*, *reliable covariation*, and *nonspuriousness*. These conditions can be met in either longitudinal or cross-sectional studies, but there is an important difference between the two. Longitudinal studies compare two or more time periods on a set of cases. The resulting regression coefficient describes the amount of *change* in the DV resulting from a unit change in the IV. Cross-sectional studies compare a set of individuals (persons or groups) who differ on the DV. The resulting regression coefficient describes the amount of *difference* between individuals on the DV, given a unit difference in the IV. Longitudinal and cross-sectional designs will yield the same regression coefficients only if the process generating variation in the DV is constant across individuals and stable across time periods.

Temporal Ordering

Although the term *cross-sectional* would seem to imply that the variables in the hypothesis are measured at exactly the same time, this is not always the case. Frequently, a study is considered cross-sectional as long as none of the variables is measured at two or more points in time—which would make the analysis *longitudinal.* For example, prediction of college freshman grade point average (GPA) from Scholastic Aptitude Test (SAT) scores would be considered cross-sectional even if SAT scores were taken 18 months before the GPAs.

Reliable Covariation

Establishing reliable covariation requires three conditions—a measure of association between two variables, statistical significance, and sample representativeness.

Measures of Association

There are many measures of association between pairs of variables ("bivariate correlation") that depend on the ways in which the variables are measured. The most widely recognized, the Pearson product-moment correlation coefficient (r), is used when both variables are continuous (either interval or ratio). Other correlation coefficients arise when the IV and DV are other types of scales (e.g., the phi coefficient is used when both variables are dichotomies) or combinations of scales (e.g., the point-biserial correlation is used when one variable is continuous and the other is a dichotomy). However, the textbook formulas for these coefficients are simply special cases of the Pearson r.

One important consideration in estimating the degree of association between two variables is whether the obtained correlation coefficient has been attenuated by unreliability in the variables or diminished by variance restriction. The maximum observed correlation

between X_1 and Y is limited by the reliabilities of the two variables. Thus, it is extremely important to measure variables as reliably as possible. This can be achieved by rationally constructing multi-item scales for both variables, pretesting these scales, conducting item analyses to refine the scales, and reporting the scales' reliability coefficients (e.g., Cronbach's α) in the study results. Variance restriction occurs when a floor or ceiling effect produces a skewed distribution of very low or very high scores, respectively. A very high proportion of low scores results when using an ability test that is too difficult or an attitude inventory whose items are too unpopular. By contrast, a very high proportion of high scores results from using an ability test that is too easy or an attitude inventory whose items are too popular.

Statistical Significance

The need to establish statistical significance arises when data are collected by drawing a sample of cases from a larger population. In such cases, the degree of association measured by the sample statistic r might differ from the corresponding degree of association measured by the population parameter ρ. The average difference between r and ρ decreases as the size of the sample increases, and procedures for statistical significance testing are used to calculate confidence intervals for the population value.

As a practical matter, sample size is important because the *statistical power* of a research design—the probability of correctly detecting a statistically significant relationship that truly exists in the population—decreases as the sample size decreases. Specification of a Type I error rate (α), desired statistical power (π), and an expected population correlation coefficient (ρ) make it possible to determine the sample size (n) needed for a given study. It is important to recognize that the resulting n is the number of respondents who provide enough data to calculate the correlation coefficient(s). As noted in the discussion of sampling below, this number must be adjusted for nonrespondents and incomplete data. That is, the sample size must be 500 if 250 are needed and there is likely to be only a 50% response rate.

Moreover, survey researchers typically find that respondents answer some items but not others. In some cases, the proportion of respondents who answer all items can be as low as 20%, so *listwise deletion* (removing all respondents with incomplete data) can severely deplete an already small sample size even though it yields reasonable estimates of standard errors for regression coefficients, b_{Yi}. On the other hand, *pairwise deletion* (computing the correlation between each pair of variables using only the cases responding to those variables) yields larger samples but can produce correlation matrixes that do not meet the mathematical requirements for multiple regression analysis (i.e., they have negative determinants and an undefined n to use in estimates of standard errors). *Unconditional mean imputation* (replacing missing values on a variable with the mean across all cases for that variable) avoids some of these problems but produces biased estimates of variances and covariances. *Conditional mean imputation* (regressing the available cases for one IV onto the remaining IVs and then using the resulting regression coefficients to estimate the missing values) is better than pairwise deletion and unconditional mean imputation but still underestimates the standard errors of the b_{Yi}. However, *maximum likelihood* and *multiple imputation* methods are likely to produce reasonably good estimates of the b_{Yi}, standard errors, and test statistics if data are *missing completely at random* (*MCAR*, the absence of data on Y is unrelated to Y or to any of the X_i), or *missing at random* (*MAR*, the absence of data on Y is unrelated to Y after controlling for the X_i). Data that are *not missing at random* (*NMAR*, there is a systematic structure to the absence of data) require even more sophisticated methods.

Sample size also is important as it relates to the number of IVs used in a regression analysis. Tests of significance for regression coefficients have only $n - k - 1$ degrees of freedom (where k is the number of IVs). Thus, as k approaches n, the estimated standard errors of the regression coefficients become increasingly large and the multiple correlation (*Multiple R*) of the IVs with the DV approaches 1.0. This problem, known as *overfitting*, will be obvious when $n/k = 1$ because $R = 1.0$ is absurd. However, overfitting will still be present, but difficult to detect, when n/k is only slightly larger than 1. Researchers are well advised to generate a sample sufficiently large that $n/k > 5$. It is also advisable to correct for *shrinkage* by reporting an *adjusted R*, which compensates for the number of predictors (see Equation 1). Alternatively, one can *cross-validate* by applying the weights estimated from one subset of the data (the estimation sample) to predict scores in another subset (the holdout sample).

$$R^2 = 1 - (1 - R^2)\frac{n-1}{n-k-1} \qquad (1)$$

Sample Representativeness

The need for sample representativeness arises from a desire to avoid distortions in the estimates of population parameters. Sample representativeness is usually sought as a way to ensure that the sample estimate of a mean (M) or proportion (p) is unbiased. A sample can be assumed to be representative if it is randomly selected from the population of interest and there is a 100% response rate. *Random selection* means that every member of the population of interest has an equal probability of being included. This is relatively easy to accomplish if an adequate *sample frame* exhaustively lists all members of the population. For example, a roster of all students currently enrolled within a given school would be an adequate sample frame if a researcher intended to generalize the results of the study only to that school. It would not be an adequate sample frame if the study results were also intended to be generalized to students in other schools or to students who might be enrolled in future years.

These more complex situations may require more sophisticated strategies for sampling a much broader population (all students in a school district, state, or the entire country). For example, stratified sampling divides all sample elements, usually individuals, into strata (categories) based upon their known attributes so that each element is in one and only one stratum. Then, a random sample is drawn from each stratum. *Proportionate stratified* sampling selects elements so that the proportion of each stratum in the sample is the same as its proportion in the population. This strategy is useful in ensuring that relatively rarely occurring sample elements (e.g., households of Native American ancestry) are represented in the sample. In some cases, researchers want to compare these strata on other variables, such as their average income, but proportionate stratified sampling would not provide an adequate sample size for the infrequent strata. *Nonproportionate stratified* sampling solves this problem by oversampling the infrequent strata—selecting a higher proportion of sampling elements from those strata than their proportion in the population. *Cluster sampling* is used when there is no sample frame from which to select a sample, so elements are selected in groups. For example, a researcher might randomly select five states from the entire country, then five counties within each of those five states, a single high school from each county, and a simple random sample of students from each high school. Cluster sampling makes it possible to draw

a nationally representative sample, but its sampling errors are greater than those for simple random sampling because the elements within a cluster are more similar than randomly selected elements from the population. This error decreases as the number of clusters increases (but, of course, increasing the number of clusters increases survey costs) and as the homogeneity of cases in each cluster increases. Although superficially similar, stratified sampling differs from cluster sampling in that a stratified sample of, for example, counties within a state would further select elements from all strata, whereas a cluster sample would further select elements only from the selected clusters.

A 100% response rate is quite rare, even in highly controlled settings such as schools and other organizations. However, there are many procedures for ensuring a high response rate. In-person interviews generally have higher response rates than telephone interviews, which, in turn, generally have higher response rates than mail questionnaires. Unfortunately, response rates for all data collection procedures have been declining significantly in recent years. In mail surveys, it is extremely helpful to keep the questionnaire short (approximately two to three pages printed in booklet form). It also helps to send a pre-letter, the initial questionnaire, a reminder postcard, and two replacement questionnaires at approximately 2-week intervals.

Because random selection and 100% response rarely occur, it is common to collect data on the demographic characteristics of the respondents and compare these to the characteristics of the population. This is feasible for general public surveys within political jurisdictions because census data on demographic characteristics such as gender, age, education, ethnicity, and income can be used to characterize the population. Although this strategy can assess the representativeness of the respondents with respect to the measured demographic characteristics, it cannot ensure that the respondents are representative with respect to any other variables. Thus, if the variables of interest (achievement motivation or fatalism, for example) are minimally related to the available demographic characteristics, it will not be possible to determine if the respondents are representative of the population with respect to the variables of interest. In sum, representativeness of the respondents with respect to the measured demographic characteristics can increase the researcher's confidence that the sample is representative with respect to the variables of interest, but cannot provide complete assurance that this is the case.

As it turns out, the issue of sample representativeness is less of an issue than it might appear. Recent research suggests that well-executed samples using Dillman's procedures tend to have respondents missing randomly rather than systematically. It is also important to recognize a critical distinction between applications such as public opinion polling and tests of causal hypotheses. The former usually report means or proportions, whereas the latter involve estimates of association (or equivalently, difference between means). Sample bias is much more likely to affect means and proportions than measures of association because the latter are mostly affected if sample bias produces a sample variance that is systematically different from that of the population. As noted above, the sample variance can be systematically restricted by a ceiling or floor effect, and this will cause downward bias in any estimated correlations. Alternatively, the sample variance can be systematically inflated if extreme groups have been selected, and this will cause upward bias in any estimated correlations.

However, there are cases in which explicit selection eliminates data on cases below (or above) a deterministic threshold. In other cases, incidental selection tends to eliminate only the highest or lowest cases. Such selection might occur on either the IVs or DVs. There are some models for analyzing *censored* samples, which have data on the IVs but not the DV. There are fewer models for analyzing *truncated* samples, in which both IVs and DVs are missing.

Nonspuriousness

A relationship can be considered spurious if the true relationship between two variables has been obscured by another variable. As the discussion of unreliability and variance restriction indicates, it is also possible for a correlation to be spuriously low. However, the most common concern in causal analysis is that the degree of association between a hypothesized causal variable (X_1) and its consequent (Y) is spuriously high. The principal ways in which an apparent causal relationship can prove to be spurious are often described as *plausible rival hypotheses.* These are illustrated in Figure 1, which presumes that a statistically significant correlation exists between X_1 and Y that is caused by the hypothesized relationship, X_1 causes Y. If this hypothesis is true, the correlation between X_1 and Y is generated by the regression of Y
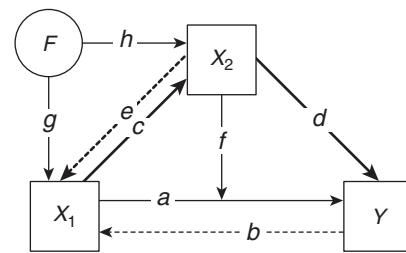
Figure 1   Basic Path Diagram

onto X_1. This is represented by path a, whose regression coefficient is b_{Y1}.

The first rival hypothesis is *reverse causation*, in which Y causes X_1. In this case, the correlation between X_1 and Y is due to path b, which is estimated by the regression of X_1 onto Y, b_{1Y}. The reverse causation hypothesis is ruled out if X_1 occurred before Y.

The second rival hypothesis is *reciprocal causation*, in which X_1 causes Y and Y causes X_1. In this case, the correlation between X_1 and Y is due to a combination of paths a and b. That is, both b_{Y1} and b_{1Y} are statistically significant. This model is usually tested using longitudinal data, but it can be tested with cross-sectional data under some conditions by using structural equation models.

The remaining rival hypotheses can best be understood by referring to Equation 2, which shows that the standardized regression coefficient for X_1 (b_{Y1}) is the same as r_{1Y} only if $r_{12} = 0$ or $r_{2Y} = 0$ (or both). Conversely, if r_{1Y} and $r_{12} \neq 0$, there is a possibility that r_{1Y} is spurious. This leads to four specific rival hypotheses.

$$b_{Yi} = \frac{r_{Yi} - r_{Yj}r_{ij}}{\sqrt{(1 - r_{ij}^2)}} \qquad (2)$$

The third rival hypothesis is *complete mediation*, in which X_1 causes X_2 and X_2 causes Y. In this case, r_{1Y} is due to paths c and d. That is, both b_{21} and b_{Y2} are statistically significant, but b_{Y1} is not. This hypothesis can be tested by regressing Y onto X_1 and X_2. If X_2 completely mediates the relationship between X_1 and Y (that is, X_2 is the direct cause of Y), b_{Y2} is statistically significant and b_{Y1} is not. Of course, b_{21} must also be statistically significant to establish mediation. If not, X_2 is an independent cause of Y that, according to Equation 2, cannot distort r_{1Y}.

A special case of complete mediation is *partial mediation*, in which X_1 has a direct effect on Y, as well as an indirect effect in which X_1 causes X_2, and, in turn, X_2 causes Y. In this case, r_{1Y} is due to paths a, c, and d. That is, b_{21}, b_{Y2}, and b_{Y1} are statistically significant. As is the case with complete mediation, a partial mediation hypothesis is also tested by regressing Y onto X_1 and X_2. Unlike the case of complete mediation, b_{Y2} and b_{Y1} are statistically significant.

The fourth rival hypothesis is *joint causation by a third variable*, in which X_2 causes both X_1 and Y. In this case, r_{1Y} is due to paths d and e. That is, b_{12} and b_{Y2} are statistically significant, but b_{Y1} is not. This hypothesis is identical to the complete mediation hypothesis except that X_2 causes X_1 rather than the reverse. Consequently, joint causation is tested in the same way as complete mediation. Joint causation can be distinguished from complete mediation if X_2 temporally precedes X_1.

A special case of joint causation is *partial causation by a third variable*, in which X_2 causes both X_1 and Y, but X_1 also causes Y. In this case, r_{1Y} is due to paths a, d, and e, so b_{12}, b_{Y2}, and b_{Y1} are statistically significant. This hypothesis is identical to partial mediation except that X_2 causes X_1 rather than the reverse. Consequently, partial causation is tested in the same way as partial mediation. As is the case with joint causation, partial causation by a third variable is distinguished from partial mediation if X_2 temporally precedes X_1.

The fifth rival hypothesis is *complete moderation*, which occurs if the regression of Y onto X_1 depends on the value of X_2. In this case, r_{1Y} depends on paths a and f. That is, $b_{(12)}$, where the subscript (12) denotes the product of X_1 and X_2, is statistically significant but b_{Y1} and b_{Y2} are not. This hypothesis is tested by regressing Y onto X_1, X_2, and X_1X_2 (the product of X_1 and X_2). If X_2 completely moderates the relationship between X_1 and Y, $b_{(12)}$ is statistically significant but b_{Y1} and b_{Y2} are not.

A special case of complete moderation is *partial moderation*, which occurs if the relationship between X_1 and Y is both multiplicative and additive. In this case, r_{1Y} depends on paths a, d, and f. This hypothesis can be tested in the same way as complete moderation. If X_2 partially moderates the relationship between X_1 and Y, $b_{(12)}$, b_{Y1}, and b_{Y2} are all significant.

The last rival hypothesis is that X_1 and X_2 are indicators of a single underlying factor F, so r_{12} is generated by the paths marked g and h. If this is the case, X_1 and X_2 can be combined by standardizing them

[i.e., $x_i = (X_i - M_i)/SD_i$, where M_i is a variable's mean and SD_i is its standard deviation] and then adding the standardized values (i.e., $F = x_1 + x_2$) to form an equal weighted composite. This hypothesis tested by regressing Y onto F and comparing the *multiple* R^2 of the equal weighted composite (which forces $b_{Y1} = b_{Y2}$) to the *Multiple* R^2 of the independent effects model regressing Y onto X_1 and X_2 (which allows b_{Y1} and b_{Y2} to differ).

Tests of the plausible rival hypotheses require that a measured third variable, X_2, be analyzed in order to estimate its influence. In some cases, there might be suspicion that an unmeasured third variable has influenced the observed correlation between X_1 and Y. In particular, some researchers have proposed that variables collected from the same source at the same time are contaminated by *common method variance* (CMV). The effects of CMV can be estimated using structural equation models or by using a theoretically unrelated variable that is likely to be contaminated by CMV but theoretically unrelated to any of the X_i or Y.

Aggregation

Researchers must recognize the distinction between the units of observation and the units of analysis. The *unit of observation* is the type of entity from which the data are collected; in most social and behavioral research, the unit of observation is the individual. The *unit of analysis* is the type of entity at which the measures are calculated and analyzed; sometimes, this is the individual, but other times, it is the group.

It is essential to be clear about the unit of analysis because that is the unit at which causal inferences must be made. It is an ecological fallacy to analyze data at the group level and draw inferences about individuals, whereas it is a reductionist fallacy to analyze data at the individual level and draw inferences about groups. Figure 2 shows how the direction of effect can be completely reversed when the proper unit of analysis is not recognized. Panel *a* shows zero correlations between X and Y when calculated at the individual level, as represented by scatterplots of the data that take the form of circles. However, systematic differences between the groups on both of the variables produce a positive correlation at the aggregate level. Conversely, Panel *b* shows negative correlations between X and Y when calculated at the individual level, as represented by scatterplots of the data that take the form of negatively sloped ellipses. However, systematic differences between the
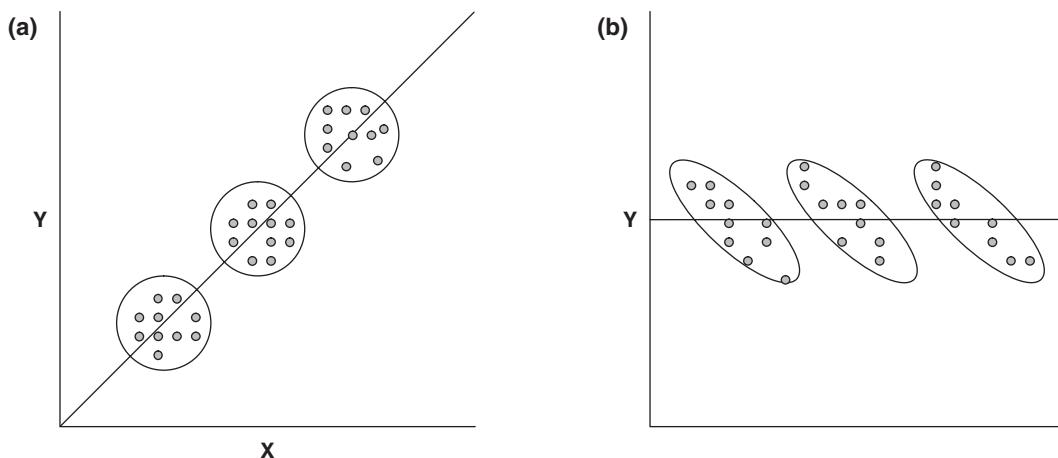
Figure 2   Aggregation Biases

groups on X, but not Y, produce zero correlation at the aggregate level.

Resolution of the disparity between results at the individual and group levels depends on which is the logical unit of analysis for the researcher's theoretical perspective. When the unit of analysis is the group, researchers often use the group mean as the score for the group. However, it will often prove useful to analyze the group's dispersion (e.g., variance) as well. To examine both individual and group processes concurrently, researchers should use hierarchical linear modeling.

Common Research Designs

The discussion to this point has assumed that all relevant variables (i.e., all variables having significant causal effects on Y) have been measured. However, this is not always the case. Indeed, common research designs are defined by the ways in which the relevant variables are treated. In addition to being measured, they can be manipulated, controlled (held constant), randomized, or omitted.

Measured variables typically involve the collection of reports made by the respondents themselves or by external observers. Such measures typically include biographical data; cognitive, psychomotor, and physical tests; personality inventories and attitude scales; and assessments of social interaction. *Manipulated variables* involve differential treatment of research participants, as when one group of students is given

an advance organizer and another is not. *Controlled variables* are held constant, as when all participants in an experiment have been selected to have the same personal characteristics (e.g., age, gender, education) or to experience the same environmental conditions (e.g., light, heat, noise, and time pressure). *Randomized variables* are ones whose effect has been probabilistically neutralized by randomly assigning participants to conditions. Random assignment makes it statistically improbable that any unmeasured characteristics of the participants are correlated with the experimental treatments. *Omitted variables* are, as the name suggests, excluded from the analysis by failing to measure, manipulate, control, or randomize them. As a practical matter, every study must omit many variables because it would be impractical to measure, manipulate, control, or randomize all of them. Thus, theory is essential in determining which variables can be omitted safely in a given study because their absence is unlikely to bias the study's conclusions. In practice, this means researchers should carefully examine the previous research to identify relevant variables and be prepared to explain which ones can be safely omitted. This means being able to document that any omitted variables that are correlated with the DV are not also correlated with the IVs (these are independent causes that will not bias the estimates of the included variables). In addition, any omitted variables that are correlated with the IVs are not also correlated with the DV (these are not the DV's other causes).

This typology of measured, manipulated, controlled, and randomized variables leads to the three principal types of cross-sectional research designs—experimental, quasi-experimental, and nonexperimental. Experimental designs have at least one DV that is measured (the presence of multiple DVs does not make the design longitudinal unless these are repeated measurements over time), and the DV need not be continuous because there are analytic procedures for analyzing dichotomous DVs. At least one IV is manipulated, and participants' assignment to treatment is randomized. Many of the other variables (respondent characteristics such as age and gender or environmental characteristics such as light and noise) are controlled in laboratory experiments, but they must be ignored in field experiments. Experimental designs usually produce the strongest causal inferences, but it is possible for a confounded design to simultaneously manipulate two distinct variables in a way that does not allow any effect to be attributed unambiguously to either variable. Experimental designs are sometimes strengthened by measuring relevant causal variables that are not controlled or part of the treatment. These measured variables are incorporated into an analysis of covariance.

Quasi-experimental designs, like experimental designs, have at least one DV that is measured and at least one IV that is manipulated. However, participants' assignment to treatment is *not* randomized. In many cases, the "manipulated treatment" in a quasi-experimental design is a naturally occurring difference between intact groups. Thus, any apparent difference associated with the treatment might be due to differences between the groups in terms of their environmental conditions or members' characteristics.

Finally, nonexperimental designs are ones in which all variables are measured. The lack of experimental control or random assignment can make it difficult to rule out the plausible rivals to any hypothesized causal relationship. This makes them weak in internal validity. However, nonexperimental designs that rely on highly credible samples, as scored by the indicators in Table 1, have high generalizability. That is, their results are more likely to be replicated in subsequent studies. In addition, they often include a much larger number of IVs, which allows researchers to apply statistical controls to variables that cannot be physically controlled. The large number of IVs and DVs in such a study can be presented in a correlation matrix that reports each variable's mean, standard deviation, and reliability

Table 1   Sample Credibility Scale

Characteristics	Score
A. Generalizability	
1. Use of geographic samples	
• Single location	0
• Several locations combined	2
• Several locations compared/limited geography	4
• Several locations compared/widespread geography	7
• The entire population or a representative sample from it	10
2. Use of special populations	
• A single sample whose bias is likely to affect study conclusions	− 5
• A single convenience sample with no apparent bias	0
• Combined convenience samples with no apparent bias	2
• Compared convenience samples with no apparent bias	4
• Compared samples from the range of population variation	7
• The entire population or a representative sample from it	10
B. Sample size	
• Sample too small to test any study objectives	0
• Sample adequate to test some, but not all, study objectives	5
• Sample adequate to test all study objectives	10
C. Discussion of sample limitations	
• No discussion of sample limitations	0
• Discussion identifies types of sample limitations	5
• Discussion assesses whether sample limitations would affect study conclusions	10
Maximum total points possible	*40*

Source: Adapted from Sudman, S. (1983). Applied sampling. In P. H. Rossi, J. D. Wright, & A. B. Anderson (Eds.), *Handbook of survey research* (pp. 145–194). San Diego, CA: Academic Press.
Note: Study designs usually involve *either* geographic samples *or* special populations but can involve both.

along with its intercorrelations with the remaining variables. It is usually helpful to list the variables in their presumed causal order—typically beginning with demographic characteristics, proceeding to psychological variables, and concluding with behaviors. This yields blocks of variables whose patterns of intercorrelations are easier to discern. In particular, it facilitates an examination of each variable's construct validity. This results when theoretically related variables are correlated (i.e., have *convergent validity*) and theoretically unrelated variables are uncorrelated (i.e., have *discriminant validity*).

Reporting a complete correlation matrix (including variable means, standard deviations, and reliabilities) allows readers to examine the plausibility of rival hypotheses that might not have been considered by the researchers or reviewers prior to publication. Such post hoc analyses are difficult, if not impossible, when researchers report only regression coefficients or, only slightly more informatively, regression coefficients and the corresponding correlation coefficients. In addition, reporting a complete correlation matrix facilitates meta-analysis, which cumulates the findings of multiple studies.

Michael K. Lindell

See also Correlation; Experimental Design; External Validity; Field Experiments; Internal Validity; Longitudinal Research; Quantitative Research Methods; Random Sample; Regression; Reliability

Further Readings

Allison, P. D. (2002). *Missing data.* Thousand Oaks, CA: Sage.

Berry, W. D. (1993). *Understanding regression assumptions.* Thousand Oaks, CA: Sage.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.

Dillman, D. A. (2007). *Mail and internet surveys* (2nd ed.). New York: Wiley.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York: Guilford.

Murphy, K. R., & Myors, B. (2004). *Statistical power analysis: A simple and general model for traditional and modern hypothesis tests* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw Hill.

Scheaffer, R. L., Mendenhall, W., & Ott, R. L. (1996). *Elementary survey samplings* (5th ed.). Pacific Grove, CA: Duxbury.

CRYSTALLIZED INTELLIGENCE

Although Charles Spearman's theory of general intelligence (g) garnered a nonnegligible amount of support during the early stages of modern intelligence testing, an accumulation of empirical research over the course of the 1930s prompted R. B. Cattell to contend that g was not unitary. Rather, Cattell argued, g was more accurately described as fractionated into two factors: crystallized intelligence (g_c), or the ability to reproduce information stored in long-term memory, and fluid intelligence (g_f). Since Cattell's proposal, an impressive amount of psychometric, neurophysiological, and individual differences research has continued to support the conceptualization of g_c, although the issue of whether the g factor is or is not unitary remains an unresolved issue. The abandonment of a strict theory of general intelligence and, in its place, the advent of theories of multiple intelligence, of which g_c invariably plays an integral part, has led to the comprehensive assessment of intelligence in children and adults as applied in practice today.

Psychometric Evidence

The psychometric or statistical evidence in favor of the validity of g_c as distinct from g rests largely upon empirical results that suggested the inadequacy of Spearman's tetrad equation for the verification of g. In simple terms, several researchers noted that the development of several spatially based tests of cognitive ability during the mid-1920s and early 1930s tended to correlate with others to a disproportionate degree than would have been otherwise predicted by their correlation with verbally/education based tests (i.e., g_c). If g were the only intelligence factor, the correlations would have been uniform across all subtests, within sampling variability. The plausibility of a g_c-type factor, as distinct from g, has also been suggested based on more sophisticated analyses, such as Louis Thurstone's multiple factor analysis, which was designed to allow for the extraction of more than one factor from a correlation matrix. More recently, the observation of

a g_c factor, as distinct from g or g_f, has also been supported based on confirmatory factor analyses.

Physiological Evidence

Cattell contended that, although psychometric evidence had a place in supporting the plausibility of g_c, the neurophysiologically related evidence should be regarded as more crucial to its establishment. Foremost, there was evidence to suggest that the trajectory of cognitive ability adult development in g_c is unique among cognitive ability factors in that performance tends to increase during early adulthood (20–30) and remains stable until the age of approximately 65, after which some decrement in performance is observed. Cattell also hypothesized that individual differences in g_c would be more greatly influenced by the environment, in comparison to other cognitive ability factors. This hypothesis has been supported by behavior genetic studies, which have suggested that approximately 45% of g_c variance may be heritable, which can be contrasted to the 65% heritability estimate associated with g_f. Relatedly, performance on g_c-like tests tends to be relatively resistant to brain injury and disease. Consequently, g_c-type tests are referred to as *hold tests* in the area of neuropsychology and frequently form the foundation for an estimate of premorbid IQ in many neuropsychological assessments. Although relatively resistant to brain damage, g_c performance has been empirically implicated to be mediated by Broca's area, which is an area of the left hemisphere of the brain that does not appear to need to be intact to perform other cognitive ability tasks. Also, in many neuropsychological cases, recovery in performance on g_c-type tests can be expected in those who have attained physical maturity. This last empirical observation supported Cattell's investment theory of intelligence, which posits that individual differences in g_c are due to individual differences in the previous application of g_f. For this reason, a positive correlation of approximately .50 is often observed between g_c and g_f. Thus, in combination with individual differences in g_f capacity, individual differences in g_c were considered to arise because of nonintellective factors such as motivation, socioeconomic status, education, and personality.

Personality and g_c

Although little research has specifically examined motivational processes that may affect performance on g_c tests, some research has examined personality dimensions as predictors of intelligence. The more recent research has focused on the personality dimension of Openness to Experience (a.k.a. Intellect and Culture). Theoretically, it would be expected that those motivated or naturally receptive to novel experiences and ideas may accumulate more knowledge than those who are less naturally disposed to such experiences. Earlier research based on bivariate correlations reported effects in the range of .25 to .30 between measures of Openness and g_c-type measures such as Vocabulary and General Information. More recently, however, the hypothesis has been tested based on structural equation models specifically constructed to partition g_c latent variable variance as distinct from g factor variance, as the two sources of variance are confounded with any measure of g_c. The structural equation model research suggests that the association between Openness to Experience and intelligence is due to g rather than g_c. Overall, the effects tend to be small, and there is some suggestion that not all elements of Openness to Experience may be associated with intelligence. In fact, certain elements of Openness, such as Openness to Feelings, may be correlated negatively with intelligence. Further refinement of personality measurement may be needed before this area of research may be expected to progress.

The Measurement of g_c: Compliments and Criticisms

Crystallized intelligence tests are perhaps the group of cognitive ability tests that tends to be associated with the best psychometric properties. With respect to internal consistency reliability, tests such as Vocabulary and General Information are very typically in the high .90s. Relatedly, the factorial nature of the items associated with typical g_c tests appear to be invariably unidimensional, an assertion that cannot be made justifiably about typical g_f tests (e.g., Raven's). Furthermore, the measurement of g_c has proven to be amenable in both oral forms and group-based/multiple-choice formats.

Scores derived from measures of g_c have, however, been criticized for being heavily confounded by social variables such as educational opportunity and socioeconomic status. They have also been described as culture specific, as commonly administered g_c tests tend to be based on one particular language. These

criticisms provided a nonnegligible amount of the impetus for the development of "culture-free" tests such as Cattell's Culture Fair test. However, some commentators have argued that typical g_c tests may not necessarily be as biased as commonly considered.

Consider a vocabulary (word knowledge) test, which may be considered one of the most valid g_c tests. A typical vocabulary test consists of several words that can be discriminated on at least two psychometric bases: (1) item difficulty and (2) item reliability. Item difficulty in the context of a vocabulary test is predicated upon the assessment of the frequency with which particular words in a language are found to appear in commonly available printed material (newspapers, magazines, books). Item reliability may be assessed by calculating item-total correlations for each item. It has been argued that g_c items that were biased by environmental opportunity would evidence low item-total correlations. Consequently, test developers can (and do) exclude such items from their published inventories of intelligence.

It has been contended that g_c and verbal intelligence are effectively two terms that refer to the same construct. However, this assertion should be viewed as, at least, theoretically untenable, as verbally based items of intelligence may be designed to be eductive (i.e., g_f) or reproductive (g_c) in nature. The critical distinction between two verbally based items that are respectively eductive and reproductive in nature pertains to whether the words that make up the item should reasonably be assumed to be known by the population for which the test has been designed. For example, the "analogies" item, "Big is to small as tall is to ____" contains words and a solution word that are extremely high in frequency usage in the English language. To solve the problem requires almost exclusively the capacity to educe relations. In contrast, the analogies item, "The Stranger is to Camus as The Metamorphosis is to ____" contains information that is not found in English-language usage with high frequency. Consequently, the item requires a nonnegligible amount of g_c to solve. Thus, although both items are verbal in nature, the first item does not measure g_c.

As noted above, g_c tests within popular intelligence batteries tend to be exclusively verbally based. However, spatially based tasks conceivably could be devised to measure g_c. Although a spatially based crystallized intelligence task may not be expected to be any less culturally biased than a verbally based g_c test, it may be argued to be more justifiable to measure g_c via more than one mode of communication or sense modality.

Factorial Status of g_c

A substantial amount of empirical research has attempted to quantify the association between g_c and g in comparison to other lower-order factors. Within the three-stratum model of intelligence, which consists of eight second-stratum factors, g_c is placed as the second most closely associated second-stratum factor to the third-stratum g factor. Given that g_c is often considered nonnegligibly influenced by the environment, a strong association between g_c and g may seem incongruent. Various researchers have attempted to reconcile this apparent contradiction by suggesting that the accumulation of word knowledge over time is not largely based upon specific instruction from parent/teacher to child/student (i.e., rote learning). Instead, the accumulation of word knowledge is based on a process of discerning or inducing the meaning of a word based on the context in which it is used—for example, based on the meaning of other words used in the sentence/paragraph, and/or based on the meaning of portions of the unknown word that are similar to other, previously learned words. Thus, the process of accumulating word knowledge would be expected to be nonnegligibly based on an eductive process, hence the strong association between g_c and g within higher-order models of intelligence. For this reason, it has been argued that no subtest should be considered a pure measure of g_c, as any valid g_c subtests would be expected to consist of both g_c- and g-related variance.

Gilles E. Gignac

See also Intelligence Quotient (IQ); Intelligence Tests

Further Readings

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies.* Cambridge, UK: Cambridge University Press.

Cattell, R. B. (1987). *Intelligence: Its structure, growth, and action.* Amsterdam: Elsevier.

Gignac, G. E. (2005). Openness to experience, general intelligence and crystallized intelligence: A methodological extension. *Intelligence, 33,* 161–167.

Mackintosh, N. J. (1998). *IQ and human intelligence.*
    Oxford, UK: Oxford University Press.

CULTURAL DEFICIT MODEL

The *cultural deficit model* (hereafter referred to as the deficit model) is the perspective that minority group members are different because their culture is deficient in important ways from the dominant majority group. The field of educational psychology has long been interested in understanding why racially different, non-White children perform differently in school, with an emphasis on academic underachievement. The deficit model has been important in the evolution of thinking about this important social issue. Hence, the deficit model asserts that racial/ethnic minority groups do not achieve as well as their White majority peers in school and life because their family culture is dysfunctional and lacking important characteristics compared to the White American culture. Other names for the deficit model have been *cultural disadvantage*, *cultural underclass*, *cultural poverty*, *culturally deprived*, and *social pathology.* In the following paragraphs, the deficit model is discussed further along with examples from schools and education.

Evolution of Thinking About Difference

Historically, U.S. researchers from various disciplines have been interested in racial differences and explaining why they exist. African Americans have been the primary focus of these early efforts since the mid-to-late 1800s. The prevailing view during these times was the *biological deficit* or *genetically deficient model:* Racial differences were caused by insufficient internal abilities, such as poor intelligence and/or inferior genes. In contrast to the biological explanation, the focus shifted to sociocultural factors related to culture and poverty—the cultural deficit model.

To explain academic achievement disparities of racial/ethnic minority children, the cultural deficit model asserts that minority cultural values are dysfunctional, which is the primary reason minority children fail to achieve academically and occupationally. The deficit model further assumes that minorities do not value education as a means of social and financial mobility. Finally, because cultural values are passed on through the family, the minority family is also viewed to be dysfunctional and insufficient.

The deficit model had become very popular during the 1960s and 1970s and influenced educational and psychological theory and research, and even political views. The deficit model's advantage is that it has clear practical implications: Schools need to provide children with the cultural experiences that they are missing at home, and help families to function better. For example, federal Head Start programs were founded to help poor minority children who came from culturally deprived families and homes.

Criticisms of the deficit model are numerous. First, the deficit model is unfair to minority children and their families, focusing the blame on their culture. The deficit model is also inaccurate because it deemphasizes the powerful effects of poverty on the families, schools, and neighborhoods, which synergistically affect academic achievement and occupational attainment. It also strongly implied that White middle-class values are superior. Fourth, the deficit model became equated with pathology in which a group's cultural values, families, or lifestyles transmit the pathology. Finally, the deficit model has limitations for scholarship because it is too narrow as an explanatory model (i.e., rigidly blames the family) for the academic underachievement of poor minority children. In short, the deficit model's negative effects are that children were narrowly viewed as "deprived" and their families became "disadvantaged," "dysfunctional," and "pathological."

In response to these criticisms, the *culturally diverse* or *cultural difference model* has more recently become popular. In this view, minority children and their cultures are different yet valuable. Alternative lifestyles and belonging to more than one culture (e.g, biculturalism) are viewed as viably different from mainstream culture and lifestyles. However, even this model, if viewed narrowly, can result in negative effects (rigidly blaming the school for academic problems) and may not account for the true complexity of how children learn and develop over time within important environments (classroom, school, peer groups, families, communities).

Social and Educational Implications

At the current time, much more attention is being given to the culturally diverse perspective formally within educational psychology and education. For example, a multifaceted approach to race and education is

becoming more popular (e.g., culturally relevant teaching and pedagogy). Still, although the exact terms of the deficit model may not be used anymore, the underlying perspective and associated beliefs of the deficit model are alive and well in society and various public service disciplines. It is critical to understand these varying perspectives and beliefs regarding the culturally different because the racial makeup of the United States is continually changing.

Samuel Y. Song and Shirley Mary Pyon

See also Asian Americans; Cultural Diversity; Culture; Ethnicity and Race; Head Start; Multicultural Classrooms; Multicultural Education

Further Readings

Baca Zinn, M. (1989). Family, race, and poverty in the eighties. *Signs: Journal of Women in Culture and Society*, *14*, 856–874.

Frisby, C. L., & Reynolds, C. R. (2006). *Comprehensive handbook of multicultural school psychology.* New York: Wiley.

Persell, C. H. (1981). Genetic and cultural deficit theories: Two sides of the same racist coin. *Journal of Black Studies*, *12*, 19–37.

Pianta, R. C., & Walsh, D. J. (1996). *High-risk children in schools: Constructing sustaining relationships.* New York: Routledge.

Sue, D. W., & Sue, D. (2002). *Counseling the culturally diverse* (4th ed.). New York: Wiley.

Cultural Diversity

Cultural diversity is the modern term giving recognition and value to the multiplicity of peoples, customs, languages, heritages, and cognitive orientations woven in the rich fabric of American society. Once thought to be a melting pot, the character of the United States is now recognized as reflecting a cultural plurality, wherein particular groups maintain their distinctive character and identity with respect and pride. Cultural diversity is celebrated as a strength, because these groups contribute their knowledge perspectives, experiences, and practices toward the advancement of the nation and humankind. The value of cultural diversity is reflected in the founding of the nation as a refuge from persecution, the recognition that all are created equal, and the prohibition against discrimination reflected in the Constitution. Cultural diversity is relevant, then, to every aspect of daily life, including education and human service. A monocultural approach to professional service is not only discriminatory but can be detrimental to the client and often a waste of the resources of the professional. This entry explores the origin of cultural diversity, as well as the essential concepts that contribute to its appropriate application to the practice of educational psychology.

Origin of Diversity

The rich cultural diversity of the United States is derived from its earliest settlement in 1607, when European emigrants from England, Spain, Sweden, Holland, and France settled with the indigenous American Indians along the Atlantic coast. Africans arrived shortly thereafter in 1619 during the slave trade. Although the settlers lived as separate colonies for 177 years, until the end of the Revolutionary War, it was not always a peaceful existence. American Indians were displaced, and wars and rebellions raged among the settlers until the British and French prevailed. Although the nation today is often associated with equality, democracy, religious tolerance, and human rights for all, history documents a sinuous journey toward these ideals.

In the next century, as the nation began to attract more settlers to its expanding territories, immigrants further increased the diversity. However, the immigration laws did not always honor the democratic ideals of the founding fathers. The Naturalization Act of 1790 restricted immigration to "free White persons." Similar exclusionary laws, such as the Chinese Exclusion Act of 1882; the Immigration Act of 1907, which stemmed the flow of Mexicans; and several other laws enacted between 1917 and 1952 excluded, set quotas, or gave preference to immigrants of specific nations until the Immigration Acts of 1968 and 1976 eliminated such discrimination. Even today, the Patriot Act of 2001 permits exclusionary practices aimed at obstructing terrorism. Thus, either by default or design, the nation is today a patchwork of cultural diversity. There is no doubt that since its founding, the United States has remained culturally pluralistic. This concept acknowledges that American society comprises diverse and distinct ethnic subsocieties that maintain their group identity.

Terminology

At this juncture, it is necessary to distinguish ethnicity from culture. The anthropological notion of culture is

an abstract concept that refers to learned rules of behavior and interaction, plus the values and beliefs that underlie the overt actions of an individual or a group of similar individuals that distinguish and identify members as a part of the group. *Ethnicity* refers to the more concrete, unchangeable outward characteristics of culture such as race, national origin, or ancestry, which are the conscious focus, not so much by members of the group, but by others. Thus, culture is an internal abstraction and ethnicity is an other-directed categorization.

Although the term *ethnicity* is popularly associated with minority groups, and sometimes religious or language groups, the further distinction of these terms is necessary. A minority group is typically defined in relation to its numerical proportion in the general population together with subtle implications related to its history, power, and dominance. However, although a numerical majority, women are sometimes categorized as a minority group particularly with regard to labor, political influence, and economic power. Specific religious groups are perceived as ethnic groups, such as Jews and the Amish people popularly known as Pennsylvania Dutch. Similar to Jews, the Amish are not "ethnically" distinct. However, their religious and cultural practices distinguish them from mainstream society. These cases exemplify the complexity of conceptually separating religion and ethnicity for some groups.

In some cases, nationality corresponds broadly to ethnicity, as in the example of Americans of recent German, Swedish, or Dutch descent. Such distinctions are readily observable, especially when national origin and language group are the same. However, a major exception is observed in the case of Spanish speakers, also known as Hispanic or Latino/a. Within this language group are subsocieties that can be linguistically, ethnically, nationally, or culturally distinct. It is important to note that not all Spanish speakers are immigrants to the United States. Historically, Spanish speakers inhabited the territories that today include Florida, Texas, California, Arizona, and New Mexico, as well as parts of Louisiana, and the midwestern and northwestern states.

Even the terms *Hispanic* and *Latino/a* reflect the diversity of cultures. *Hispanic*, often used interchangeably with *Latino/a*, is a designation adopted by the U.S. Census Bureau to classify any individual of Spanish ancestry. Contrastively, *Latino/a* is a term adopted recently by younger non-European individuals to express their cultural distinction and pride.

Inextricably associated with cultural diversity is linguistic diversity. Currently, two main languages, English and Spanish, predominate in the United States, although there are more than 100 languages represented among its bilingual citizenry. During the European wave of immigration, assimilation was favored over preservation of the original culture and language. English was valued as the lingua franca and unofficial national tongue. High importance was attributed to learning English, especially Standard English, as the required medium for identification and socioeconomic advancement. However, for particular groups, such as African Americans, Hispanics, and American Indians, language preservation is a source of cultural identity, pride, and solidarity. Currently, Spanish has gained recognition in government, social, and business interactions, and bilingualism has gained a measure of acceptance with the latest wave of immigrants since the 1990s.

The language of African Americans, known as African American English (AAE) and colloquially known as Ebonics, has been the subject of social and political concern, as well as educational research. A dialect historically thought to be the faulty learning of, or misuse of, Standard English, linguistic scholars presently acknowledge the origin of AAE as a fully developed linguistic system derived from a Creole of Western African languages together with English, Dutch, Portuguese, and other European languages spoken by the colonizers of the African continent. Contrary to popular thought, AAE was transported to the United States and the Caribbean during the Middle Passage, survived among the African slaves, and has undergone a process of decreolization, or merging with Standard English, since the abolition of slavery. Similar influences of this linguistic merging are observed in Haitian Kreyol, Afro-Brazilian Portuguese, and Afro-Cuban Spanish. Although it has been erroneously asserted that AAE affects intelligence, academic performance, and literacy acquisition, direct evidence for this claim is generally lacking, as the seeming effects are more likely due to racism; de facto segregation; and other, more complex factors.

Another prominent component of cultural diversity is cognitive orientation, most specifically sexual preference. Research has shed light on the possibility of a biological genesis for homosexuality. Certainly, most practicing homosexuals will reveal that their sexual proclivity is no more subject to their own selection than for heterosexuals. There is no argument that open homosexuals identify as a cultural group.

A final cultural group for consideration has ties across all others previously discussed. This group is known as Deaf Culture. Congenitally deaf individuals qualify as a cultural group in every sense of the word. Historically classified as disabled, congenitally deaf individuals have recently asserted their right to be recognized as fully functional members of society. A full understanding of deaf culture requires examination of the dimensions of deafness.

Deafness can be congenital or acquired. Generally, individuals with acquired deafness are classified as members of the hearing community, and they accept their acquired condition as a disability. The rights of these individuals are protected under the Americans with Disabilities Act. Much like individuals with a partial hearing loss, they usually do not identify with Deaf Culture. Unlike their congenitally deaf counterparts, they prefer oral communication for full interaction with the hearing community.

Contrastively, members of Deaf Culture prefer American Sign Language (ASL) as their medium of communication. Thus, Deaf Culture can also be classified as a linguistic group. Historically, teaching oral communication and efforts considered as attempts to correct a disability, including speech therapy and cochlear implants, have been rejected.

A major characteristic of Deaf Culture is self-determination. Much like the Amish, assimilation into general society, especially in education, is not the preference as deaf individuals prefer separate education in deaf schools wherein the culture is practiced and preserved. Cultural conflict can arise when a deaf child is born to hearing parents who promote assimilation.

Having discussed the major cultural groups and the salient issues within the context of society, this entry now examines the application of this information to professional practice using three case examples. The reader should bear in mind that the rules of intercultural interaction are dynamic, and extremely few universals apply.

Professional Issues

Case Example 1

An Islamic female student in counseling psychology is assigned to practicum with an adult male client. Privacy rights of the client require that the session be held in a closed environment. The student insists that this would violate her religion. The instructor advises the student to transfer to another field.

Cultural conflict arises when the rights guaranteed to one individual impinge upon the perceived rights of another. Certainly, the client is guaranteed privacy, yet the requirement goes against the religious practices of the student who has the similar right to receive her education. Often, these cases find their way to the judicial system with much toil and chagrin. However, in the professional setting, a sense of mutual compromise should prevail. Two scenarios represent a possible compromise.

1. The client's right to privacy would not be violated by the presence of the instructor. The instructor could be present in the room during the counseling session.

2. With the door ajar, the counseling area could be sectioned off, and access to the entire area restricted during the time of the session.

Note that in each scenario, protection of the client's rights is the priority because these rights are mandated by law.

Case Example 2

A deaf student with a cochlear implant (the reader may substitute an individual who stutters, or a speaker with a foreign accent) is accepted in the speech-language pathology program. The student is not permitted to conduct clinical practicum and is granted a "nonclinical degree." This degree will not enable the student to practice in the field.

Has the university violated the student's rights? This question is complex and remains controversial. The university will grant the student a degree, but the student will not be qualified to practice in the profession. Whether the university's action is unethical depends upon a strict or loose interpretation of the purpose of the degree. With a strict interpretation, one may argue that the degree is but one criterion for professional practice. And because there are other certification requirements, even the regular clinical degree does not guarantee entry into the profession. With a loose interpretation, one would argue that without the required practicum and the regular clinical degree at a minimum, the university itself has restricted the student from professional practice.

Generally, in cases such as this wherein discrimination in admission policies is prohibited, the burden rests with the university to provide the necessary accommodations to afford the student equal access to the degree and the profession.

Case Example 3

In compliance with a mandate to achieve diversity in the workforce, an elementary school hires African American, Hispanic, Asian, and Middle Eastern teachers. Several parents state to the principal that they desire for their children to be taught only by White teachers.

Should the principal comply with this request? Perhaps there is a temptation to comply with the parents' request because the action would not be openly observable and no one would be hurt. Moreover, parents often select their child's teacher based on other, different characteristics.

It can be argued that such subtle forms of discrimination are the most pernicious. *White privilege* is the term used to depict the covert advantage afforded to Whites that is characteristically unknown to those against whom it discriminates. An additional subtle form of discrimination relates to the perceived value of an individual to the society; for example, the death of, or an atrocity involving, a White person arouses a greater public reaction than the same for a non-White person. For example, a missing African American teenager may barely receive mention in the media, whereas a missing White teenager may be cause for an Amber Alert.

In the foregoing case example, personal and professional ethics should dictate the principal's decision not to comply with the parent's request. The ravages of subtle discrimination are borne by the profession and the society inasmuch as their cumulative effects promote cultural divisiveness and ill will.

Political Incorrectness

Several noted professionals, celebrities, and public figures have had their careers ruined by pejorative slips of the tongue regarding cultural groups. Either as an ethnic joke overheard, slang use of a name, or stereotypic generalization, political incorrectness has become intolerable. Political incorrectness is regarded as a manifestation of covert prejudice. Revelation of such prejudice renders anyone of professional position, or social or political influence, untrustworthy in their overt actions or decisions. As a case example, in response to violence at a predominantly White suburban elite high school, a White principal reassured the parents that the incident was "Black on Black" crime. Only the overwhelming support of African American parents citing previous instances of cultural fairness saved the principal's career from destruction.

Political incorrectness is dictated by the principle of "Once spoken, there can be no retraction." Consider the recent example of the political candidate who gained notoriety by referring to his mother endearingly with a pejorative slang term. Similarly, a different candidate alluded to his African American opponent as "articulate." Although covert prejudice in these two examples may not have necessarily been intended, it was the irresponsibility of the speaker that was not tolerated.

Intercultural interaction requires knowledge of insider versus outsider rules of communication and conduct. For example, it may not be taboo for the slang pejorative term for African Americans to be used by African Americans in casual communication. However, any use of the same term by a non–African American can evoke immense pain, and often a severe negative reaction. Suffice that African Americans only can sense the deep historical assault associated with the term, and ironically, insider use may somehow reduce the impact of the assault. However, this emotion is not transferable, so any use of the term by an outsider is considered intolerable.

Test Bias

A major professional issue with regard to cultural diversity is test bias. Most professions rely on standardized tests for decisions ranging from determining who requires services to measuring the competency of professionals. A standardized test is one that is scored on the basis of the performance of a norming population. Rarely is this population reflective of the numerical proportion of every cultural group. Hence, standardized tests present a bias if they unfairly overgeneralize or do not reflect appropriate sensitivity in regard to a specific cultural group. This is not to imply that a standardized test is never accurate for individuals from cultural groups. But wherever systematic results are reflected for a large number of individuals from a cultural group, a bias should be

suspected. Such biases lead to misdiagnosis and inappropriate services.

There are several possible sources of cultural bias. A major source is value bias. Value bias is reflected in test items that represent a mismatch in the cognitive assumptions of the test maker and the examinee. An example of value bias is demonstrated in the case when the examiner expects the examinee to state answers using full sentences or narrative scenarios, and the examinee provides a one-word response and presumes that the examiner can fill in the remainder. This phenomenon of "shared knowledge" is often observed in high-context/low-verbal cultural groups such as African Americans, Hispanics, and American Indians. A high-context/low-verbal culture relies on the context, environment, and nonverbal responses. Thus, fewer words are necessary to effect communication.

Situational bias is often overlooked when examining children and adults with standardized tests. Situational bias refers to the social context, format, assumptions, and implicit rules of the testing moment. Many students report that they do not perform well on multiple-choice examinations because they are not accustomed to making one correct choice from a multitude of possibilities. Also, when the test environment assumes a time-restricted or competitive nature, cultural groups such as American Indians and some Hispanics may subconsciously reject the testing exercise in its entirety. In other cases of situational bias, there may be a fallacy that the examinee will approach the test as her or his "ideal self," as characterized by the popular adage "Putting the best foot forward." A tendency of several cultural groups is to take the test at face value and respond from a deeply personal vantage. The wrong answer is incurred, not because the examinee did not possess the answer, but because the perception of the right answer reflected a cultural mismatch of expectations.

Another type of bias is known as linguistic bias. Linguistic bias may involve the grammatical constructions within the test items, wording of the instructions, or even scoring of the responses. The results of linguistic bias reflect the interference of the examinee's native language or dialect.

The use of culturally and linguistically discriminatory instruments is prohibited by law. However, compliance may be impossible wherever truly culturally fair and culturally sensitive instruments do not exist. The solution is that professionals should recognize possible biases, apply appropriate accommodations, refrain from total reliance on standardized test scores, and demonstrate the validity of test results with additional nontest evidence.

Intelligence and Cognitive Style

The debate on culture and intelligence has existed for more than a millennium. At one time, Europeans were thought to be less intelligent than Africans. In the United States, the debate emerged in the 1960s, and again in the 1980s, as Whites were observed to score higher on traditional intelligence tests than African Americans. Toward the resolution of the debate, Howard Garner first advanced the psychological theory of multiple intelligences in 1983. Gardner posited that the typical constructs used to define and measure intelligence were too narrow. To date, he has identified eight separate intelligences—linguistic, logical-mathematical, spatial, bodily-kinesthetic, musical, naturalistic, intrapersonal, and interpersonal—for which individuals variously demonstrate proclivities. Few intelligence types listed by Gardner lend themselves to academic performance.

Modern theories of the relationship of intelligence and culture espouse either the nature or nurture position. Nature theorists follow the assumption that intelligence is congenital, biological, and genetic. Certainly, congenital abnormalities such as mental retardation support their position. Nurture theorists consider the interplay of environmental factors such as nutrition, parental education, home environment, socioeconomic level, cultural beliefs, motivation, language barriers, quality of education, health, racism, and lack of positive role models as determinants of intellectual ability.

Cognitive style, also known as learning style or conceptual style, supports the notion of multiple intelligences by recognition that individuals encode and decode input information and solve problems according to their predominant tendency of thinking and reasoning. Cognitive styles have been broadly associated with various cultures, albeit without absolute reference. Rosalie Cohen is one proponent who proposes a dichotomous model of cognitive style. In this model, the Analytic learning style, which favors White, middle-class individuals, is identified by the tendency toward stimulus sensitivity, focus on detail, objectivity, and lengthened attention span. Contrastively, the Relational learning style, which characterizes low-income minority groups, features the tendency toward subjective learning, learning in a cooperative social context, personalized content, and shortened attention span.

Cohen describes how the classroom favors the Analytic style, leaving Relational students at great disadvantage. For example, teaching models used in traditional American classrooms require the use of linguistic, logical-mathematical, and interpersonal intelligences, which are associated with the Analytic cognitive style.

Achievement Gap

The *achievement gap* refers to the performance differential in academic performance wherein racial, linguistic, and cultural minorities lag behind their mainstream counterparts on such indicators as reading achievement, grade point average, and dropout rate. It is observed that by the time they reach twelfth grade, provided they do, minority students are 4 years behind in their achievement, with African American and Latino students demonstrating skills in English, math, and science similar to those of eighth-grade White students. Similarly, deaf students typically fail to achieve grade-appropriate reading levels.

It is evident from national measurements collected over the past three decades by the National Assessment of Educational Progress that there has been little change, even when the socioeconomic level is controlled. Schools have witnessed an overrepresentation of minorities in special education programs, and a similar underrepresentation in programs for the gifted and talented. It is not known how much the issue of test bias, and the overrepresentation in special education, contribute to this situation.

Although it is generally accepted that the achievement disparities for African Americans and Latinos reflect the confluence of a multitude of variables, including language, income, family structure, schooling, and cultural factors, these factors alone do not adequately explain the genesis of the gap. Hence, school reform has been a top priority since the mid-1980s. This movement has enjoyed many successes, but significant challenges remain.

Health Disparities

The inferior health status of minority populations in the United States is well documented. Compared to Whites, African Americans, Latinos, Asians, and American Indians have a greater prevalence of diseases, higher mortality, shorter life expectancy, and poorer treatment outcomes. Diseases and chronic conditions, including cancer, hypertension, asthma, diabetes, and lead poisoning, occur with much greater frequency and are more likely to result in hospitalization and death in African Americans and Latinos. HIV/AIDS is a condition historically associated with the male homosexual community, yet new cases are increasing fastest among minorities, particularly African American women and young adults of 18–25 years.

The causes of health disparities can be categorized in four major areas: genetics; cultural, socioeconomic, and environmental factors; access to health care services; and quality of the care provided. With regard to genetics, greater emphasis is now being placed on familial predisposition to illness. Particular cancers, hypertension, obesity, and diabetes have been observed through successive generations. Cultural practices, such as unhealthy dietary intake and a sedentary lifestyle, together with genetics increase the probability of disease conditions. Particular conditions, such as fetal alcoholism syndrome in American Indians, also reflect the influence of cultural practices and genetics.

Cultural factors also include beliefs and attitudes about the body and life, as well as preference for home remedies, which are all characteristic of minority populations. For example, African Americans, Latinos, Asians, and American Indians are more likely to seek holistic or nontraditional remedies such as herbalists, faith healers, and shamans.

Socioeconomic causes of health disparities include lack of insurance coverage, inadequate insurance coverage, lack of transportation, and lack of a consistent source of health care. Minorities are more likely to use emergency and acute care as a last-resort effort, rather than pursue regular visits to a physician.

Linguistic barriers and lack of health information also contribute to health disparities for minority populations. The linguistic barrier is particularly apparent for Deaf Culture and bilingual immigrants. Immigrants also encounter legal barriers because federal law prohibits Medicaid coverage until individuals have resided in the country for 5 years. These barriers restrict minority populations' access to health care.

The burden of health disparities is not entirely the responsibility of minority care receivers. Scarcity of health workers, especially in rural areas, as well as lack of diversity in the health care workforce and lack of cultural competence among workers are also barriers related to quality of health care. The issues associated with health disparities are currently being addressed assiduously, with their complete elimination as a national goal.

Professional Service Delivery

Every culture displays its own unique qualities. But no culture is superior to another because each culture functions to preserve the well-being and highest quality of life for its members within the context of its unique worldview. Culture competency requires a knowledge of the ideas discussed in this entry, plus the acknowledgment that every individual deserves the highest level of professional service. An appreciation of cultural differences is an opportunity for self-enlightenment, and the use of culturally relative and culturally competent methods of service delivery is the key to a stronger and more effective profession.

Kay T. Payne

See also Bilingualism; Cognitive and Cultural Styles; Culture

Further Readings

Cohen, R. A. (1969). Conceptual styles, culture conflict, and nonverbal tests of intelligence. *American Anthropologist, 71*(5), 828–856.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences.* New York: Basic Books.

CULTURE

Although various attempts have been made to define *culture*, traditional definitions of culture have been very limited in scope. Historically, *culture* has been defined as a series of traits shared by a group of individuals. That is, the concept of culture has been understood as shared norms that shape individual behavior. These behaviors have been interpreted as being standardized rules for a group of individuals. Despite the unifying traits that organize and, at times, define cultural groups, culture has been traditionally construed as a monolithic variable that is stagnant or fixed both in time and space. Defining *culture* in such a way fuels a discourse that essentializes individuals of particular cultural groups by universalizing the existence of a uniform set of traits that is shared by each group. What is problematic is how this definition reinforces a perspective that universalizes and/or standardizes group norms, behaviors, and ideologies. Doing so perpetuates a discourse that reinforces these behaviors and ideologies as being static and neither fluid nor dynamic.

Emerging, however, are attempts to understand culture as the everyday practices and beliefs that individuals embody and display through behaviors and practices. This perspective reinforces discourses suggesting that individuals communicate their culture through language, rituals, religious beliefs, value systems, traditions, and other beliefs. In other words, culture has been constructed as a way of "doing life" or engaging in life-centered activities. Given this orientation, culture is a socially constructed concept that is dynamic, emergent, interactional, and multidimensional; it is not constant, fixed, or compartmentalized within a particular category or definition.

Rather than focus on what culture is and is not, the aim of this entry is to reflect on the practices in which individuals engage within a particular context and within a particular moment in time. Embedded within this perspective is an affirmation of how individuals engage in life and how individuals discuss their engagement. This entry examines the construct *culture* by moving beyond uniform categorizations of shared group cultures that tend to essentialize groups and their respective practices. The entry's purpose is threefold. First, it provides a general understanding of culture within the field of educational psychology. The complexities of culture are then illustrated by examining how research reports learning styles and the variation in learning outcomes. Second, a critical discussion is presented on the complexity that arises when using broad definitions of culture and ethnicity, especially when analyzing and discussing specific cultural groups. A discussion highlighting the strengths and weaknesses of using broad-based terms to define specific cultural groups when designing, conducting, and interpreting empirical research follows. Finally, there is a discussion on how constructs such as ethnicity and race are often used interchangeably with constructs such as culture.

Culture and Educational Psychology

Culture is a complex process that, in the broadest sense, can be defined as the means by which a group of individuals engages in daily activities over the life course. It is the commonalities such as norms, beliefs, language, and values that are shared by a group of individuals that shape culture. Moreover, culture is considered to be a dynamic and multidimensional source of influence on the developmental processes within a community. Understanding culture as a sociocultural

construct within a particular context reinforces the importance of examining both macro and micro dimensions that affect how culture is operationalized. Of significance to the field of educational psychology is how learning variations among school-age students and culture get operationalized and are understood.

Given this, recent research has advocated that individual development is not affected by culture. Rather, individuals develop as they participate in activities and/or daily practices that they construct, yet practices also include an inheritance of ideas from previous generations. This perspective centers on an ideology that positions individuals as having an impact on culture rather than culture changing the individuals. Such an approach also debunks the ideology that culture is a category into which individuals fit. It reinforces the notion that culture encompasses the cultural practices lived by individuals.

Historically, the discrepancy in academic achievement among students in the United States has been attributed to the students' culture. Research reports that the majority of students who have low academic achievement outcomes are from poor and working-class backgrounds and tend to be students of color. Suggesting culture as the cause for academic achievement differences reinforces a paradigm that places the individual's culture at the center of the problem. That is, a deficit model paradigm is used to understand the learning variabilities and outcomes evidenced among students. Of serious concern is how such a discourse frames the problem as being culturally based (i.e., an individual's culture). Such reasoning suggests that the individual's culture must change because it is the "problem," and thus educational institutions and U.S. society are absolved of any responsibility. Such a discourse also reinforces an assimilationist approach, suggesting that individuals who are not from the dominant culture must assimilate and reflect practices that are in alignment with the dominant culture. For example, it is common to hear how ineffective parenting or lack of parental interest in the child's schooling is the main cause for low-performing students. The responsibility of improving the student's learning outcomes is placed on the individual and responsibility is removed from the educational system (i.e., schools) and U.S. society. Although strides have been made to eliminate deficit models that frame academic achievement disparities, they are still salient within the U.S. educational system and evident in some policy reports.

A recent example of an approach used to overcome culturally deficit theoretical frameworks to understand learning variabilities among children and youth in U.S. schools is the cultural historical approach. Gutierrez and Rogoff argue that a

> cultural historical approach offers a way to get beyond a widespread assumption that characteristics of cultural groups are located within individuals as "carriers" of culture—an assumption that creates problems, especially as research on cultural styles of ethnic (or racial) groups is applied in schools. (p. 19)

For example, cultural styles are no longer viewed as deficits because they are examined from the community's participant's perspective. Alternative models suggest omitting the culture embodied by participants from the analysis. Such models favor paradigms that deny cultural variations, implicitly suggesting that the cultural practices of the dominant group were the norm and the developmental trajectories and learning outcomes for the dominant group (i.e., White middle class) were normalized.

The movement of educators to gain awareness and understand the culture of their students was also a promising attempt to discredit the cultural deficit approach. Of concern was how school, home, and/or community cultures juxtapose each other. Knowing this, solutions were crafted that focused on training teachers to engage in cultural patterns that were in sync with the students' cultural worlds. Implicit within this view is how cultural groups are understood as being confined by normative practices that are clustered together to form a unifying culture. Such definitions affect how educational institutions operationalize their curricula. For example, pre-service teacher education programs may train teachers to be culturally sensitive.

Culture, Race, and Ethnicity

Historical definitions of culture, suggesting that individuals from similar cultural groups display similar behaviors based on traits, perpetuate and reinforce stereotypical cultural behaviors. When defined by traits, culture suggests that the traits occur independently of people. Defining culture as a series of traits fuels a framework that embraces a deficit approach when examining learning variations among students of color. Serious implications arise when one defines culture as

a series of traits and variations are attributed to distinction in cultural traits among cultural groups. For example, this approach reinforces cultural traits as fixed—static constructs that remain constant across time and space. Recent work such as that by Banks and Gay examine cultures within a particular context and at a particular point in time, emphasizing the dynamic and more situational view of cultural practices.

Given this, individuals *live* culture. Thus, culture is a process that is dynamic, interactional, and contextual, one that is constantly changing over time. When culture is examined at the macro societal level, the intersection of race, class, and gender becomes noteworthy. Furthermore, historical and political issues become part of the analysis in understanding how culture (trans)forms over time. Recent legislation, such as No Child Left Behind, has affected the experience of children and youth in U.S. schools. That is, the high-stakes testing assessment implemented in schools has changed the learning culture for school-age children.

Similar to culture, race is also a socially constructed variable whereby the stratification of groups, community, and individuals becomes especially noteworthy. Although some researchers have argued that group labels reflect transformations in awareness in the sociopolitical milieu, these group labels are also used by the privileged in our society to impose negative attitudes and oppression onto the various ethnic working-class populations. It is through the social construction of group labels that racial formations are created and understood. These racial formations have become what we understand as ethnicity.

Ethnicity has been defined as a collective identity and is often explained in terms of national group membership and/or religious affiliations. Hence, a distinction can be made when defining and interpreting racial categories given the unique cultural affiliations that mark ethnicity within racial subgroups. More specifically, ethnicity is a critical element of race in that it is the foundation of diversity within and between racial categories. For instance, although "Hispanic American" constitutes a "racial category" in the United States, there is great ethnic diversity among Hispanic Americans that is often overlooked. For example, those of Cuban, Puerto Rican, Mexican, and Central American descent have cultural distinctions that mark how life is lived.

Given the complexity that underlies definitions of culture and race, it is critical that scholars begin to "unpackage" definitions of ethnicity as they attempt to study cultural contexts. Individuals must take notice of their position in society and articulate their marginalization, which is rooted in ethnicity, social class, migration, and citizenship. The complexity of ethnicity is revealed as individual lives are contextualized within the sociopolitical and historical experiences these individuals and their communities have endured. Acknowledging intragroup differences consequently reveals the multiplicity of experiences subsumed under the broad-based term *Latina/o*.

Complexity of Broad Definitions

Of particular importance is how researchers study cultural context. Given how culture is defined, it would be important for researchers to study cultural context in the making or as it is being lived, rather than as a cultural object or a stagnant, idealized concept. There are several strengths in using a broad definition of culture and ethnicity when studying ethnic group populations. For example, researchers can obtain a sample of participants who affiliate with the group label of Native American more readily than with specific tribal affiliations like Navajo Indians. Such general sample populations are more easily accessible when geographical location is not considered. For example, Latinas/os in California might include a higher concentration of Chicanos or recent emigrants from Mexico than, say, a population of Latinas/os in Massachusetts, where there might be a higher concentration of Cubans and Puerto Ricans. Hence, researchers who use a broad definition of culture and ethnicity and conduct research in either location would still be able to make claims about the Latina/o population despite the different subgroup affiliations. The major strengths of such studies can be generally described as follows:

1. Increased availability and accessibility of sample populations to be studied

2. More resources and information pertaining to the population being studied, because information is not always aggregated by ethnicity, social class, geographic location, immigration, and so on, but is subsumed under general racial categories such as "Latina/o" or "Native American" (i.e., census data, school and educational testing statistics, health and social service statistics)

3. Facilitation of discussion of individuals from multi-ethnic or mixed-race backgrounds because they are

included as members of specific racial groups (e.g., Latina/o, Native American).

Intragroup heterogeneity exists along varying positions such as ethnicity, social class, and regional identification. Such distinct and complex affiliations prevent ethnic or racial group membership from being equated into one common cultural experience. When one uses broad-based terms to characterize an individual and culture, one discounts the cultural diversity embedded within a particular ethnic group. The terms *Native American* or *Latina/o* will be used to illustrate this point.

The terms *Native American* and *Latina/o* have been socially constructed to include individuals from either Native ancestry and/or Latina/o ancestry. Despite the many core traditional values that Native Americans share across tribal groups, it should be recognized that Native Americans are a heterogeneous group. Native Americans differ vastly in their level of acceptance of and commitment to specific tribal values, beliefs, and practices through a variance of customs, language, and type of family structure.

The cultural differences that exist among the Native American populations are extensive, yet in some cases extremely subtle. However, it is not sufficient enough to distinguish these cultural differences by solely recognizing and acknowledging the sociocultural behaviors of individual Native American tribes. Rather, one must also understand Native American families through a lens that is sensitive to and understanding of the specific demographic variables. For example, the primary locale of residence, whether it is in an urban or reservation area, plays a critical role in Native American ethnic identification and instillation of cultural values. The families' socioeconomic status and family structure are also important variables to better understand contemporary Native American families.

The *Native American* term, albeit an umbrella term, does not capture the distinct cultural and spiritual traditions that make up each tribe of the more than 500 federally and state-recognized Indian nations and slightly more than 200 unrecognized Indian nations. For example, the Hopi Indians have substantially distinct traditions that differ from the Oneida Nation. Unfortunately, the cultural distinctions that exist between various tribes are often ignored and not accounted for when using broad terms such as *Native American.*

The term *Latina/o* is a broad ethnic label encompassing Puerto Ricans, Mexican Americans or Chicana/os, Cuban Americans, and any other individuals whose ancestry may be from Latin America. As in the case of Native Americans, the broad-based term *Latina/o* ignores historical, social, political, and cultural histories and traditions of the specific populations. More specifically, the term *Latina/o* camouflages the intragroup differences that exist (e.g., country of origin, generational history, familial immigration patterns and history, language preference, etc.). Usage of a broad-based term also ignores political struggles and historical experiences that each group has endured. Some groups of Latinas/os residing in the United States do not have the same historical and sociopolitical relationship to the United States as other Latina/o groups. For example, Chicanas/os and Puerto Ricans are two groups of U.S. citizens whose historical relationship to the United States were not originally shaped by immigration or forced exile, but rather by conquest and colonization. Unfortunately, broad-based terms such as *Hispanic* and *Latina/o* do not acknowledge these historical experiences, nor do such terms accurately characterize the individual lives and experiences of Chicanas/os and Puerto Ricans as legal citizens in the United States. Such examples illustrate the importance of acknowledging racial, ethnic, and geographical locations and family structure coupled with individual cultural factors, including migration, in an effort to better understand how such conditions influence one's development.

Another illustration of how broad-based ethnic terms may not accurately characterize ethnic groups is exemplified when researchers refer to a population as being of Mexican origin. This term refers to individuals who have recently immigrated to the United States, and it also includes individuals who have resided in the United States for generations coupled with individuals who resided in the United States when it was part of Mexico. Collapsing the experiences of individuals from various generations fails to acknowledge the distinct cultural experiences attributed by generational status in the United States.

Viewed in this way, the Latina/o population is too often solely characterized by its cultural relationship to its home or mother country. For example, it is assumed that Latinas/os migrating to the United States bring specific cultural traditions and familial practices as recognized in their home country. However, overlooked when making this distinction is how narratives

of migration are distinct and carry unique understandings that are dependent on the population, the population's proximity to its home country, and the regional differences that exist in the population's country of origin. Underlying each of the narratives will be the political climate that frames the discourse toward immigration and immigrants.

These generational, linguistic, regional, and economic differences are significant and should be accounted for when conducting research on developmental processes as they will account for multiple influences of our own cultural and sociohistorical locations. For example, Jeffrey Lewis points out how indigenous populations like the Mixtec, who live in Mexico and who migrate to the United States, form a different relationship with the United States than other recent emigrants from Mexico. The Mixtec are noted for their transnational migratory practices, spending equal amounts of time in both the United States and in Oaxaca, Mexico. This continual interaction with their home country creates a unique relationship with the United States such that they maintain their indigenous beliefs, language, and cultural practices because of the continual interaction with their native communities. It is such networks of cultural flow that maintain a specific ethnic relationship for the Mixtec, who are indigenous to their region of Mexico and not to Mexico as a whole. Unfortunately, the specificity of this cultural relationship is not acknowledged when subsumed under the term *Latina/o* or by the term *Mexican immigrant*.

Research Issues

Some scholars have begun to argue that socialization patterns differ for ethnic minority populations, reinforcing the importance of ethnic-specific research studies. Research studies that use broad-based terms are subject to misleading interpretations when the culturally specific constructs of one's ethnic background are omitted. In many cases, such studies view Euro-American middle-class families as the norm and frame developmental paradigms to fit this particular group. Conducting research in this manner distorts research findings by failing to acknowledge that European Americans are also part of our racialized society and hence part of a racialized group with a particular social position that grants privileges not accessible by other racialized groups. Thereby, broad-based terms also discount the heterogeneity within European

Americans. These examples illustrate the importance of not dismissing the fact that memberships to social, ethnic, and religious groups are complex and multidimensional, given that groups can be defined by race, gender, age, religion, social class, language, immigrant status, minority status, sexual orientation, and ethnic identification.

Scholars have pointed out that if comparative research studies are planned, it is important to assess them for cultural validity. Cultural validity focuses on the identification of rules that regulate individual and group conduct and the rules that define practices and institutions. Obtaining knowledge of individual and group rules and practices is critical to understanding a particular culture and cultural differences. Through this knowledge, researchers are able to determine the degree to which certain ideas differ in meaning across and within cultures. Incorporating a variety of research methodologies, ranging from focus groups to participant observations in community neighborhoods, family functions, and school settings, will ensure culturally valid assessments.

The interplay between culture and developmental processes is very complex and best captured using a combination of methodologies. Multiple methods of inquiry, assessments, analysis, and interpretation must be employed in order to comprehend the impact of cultural processes on adolescent development. Albeit common, the usage of a certain assessment tool or method of inquiry constructed for one population, yet employed in another population, suggests that one standard is the norm. As discussed, this assessment becomes problematic. Embedded within these methodological issues is the manner in which categories are constructed to define the population or the context being researched. For example, if a concept or object has diverse meanings among individuals of different cultures, then it is challenging to compare data on the concept or object across cultures or assume that the research collected based on these discrepancies is valid and hence advancing our current knowledge. Furthermore, scholars caution researchers to be critical of using the same measure or instrument to assess individuals within a single ethnic or racial group because of the intragroup diversity. It is important to acknowledge that these measures were designed for a particular group and may not have measurement equivalence across or within ethnic groups. Researchers who conduct comparative research studies with culturally different

groups are encouraged to evaluate the degree of interpretive validity.

Patricia D. Quijada

See also Cultural Diversity; Diversity; Ethnicity and Race; Multicultural Classrooms; Multicultural Education

Further Readings

Beauvais, F. (2000). Indian adolescence: Opportunity and challenge. In R. Montemayor, G. R. Adams, & T. P. Gullotta (Eds.), *Adolescent diversity in ethnic, economic, and cultural contexts* (pp. 110–140). Thousand Oaks, CA: Sage.

Garcia Coll, C., & Magnuson, K. (1999). Cultural influences on child development: Are we ready for a paradigm shift? In A. Masten (Ed.), *Cultural processes in child development: The Minnesota symposia on child psychology* (pp. 1–24). Mahwah, NJ: Lawrence Erlbaum.

González, N. (2005). Beyond culture: The hybridity of funds of knowledge. In N. González, L. Moll, & C. Amanti (Eds.), *Funds of knowledge: Theorizing practices in households, communities, and classrooms* (pp. 29–46). Mahwah, NJ: Lawrence Erlbaum.

Gutierrez, K. D., & Rogoff, B. (2003). Cultural ways of learning: Individual traits or repertoires of practice. *Educational Researcher, 32*(5), 19–25.

Haney López, I. F. (1995). The social construction of race. In R. Delgado (Ed.), *Critical race theory: The cutting edge* (pp. 191–203). Philadelphia: Temple University Press.

Hemant, S., & Thornton, M. C. (2004). *Newspaper coverage of interethnic conflict: Competing visions of America.* Thousand Oaks, CA: Sage.

McLoyd, V. C., Hill, N. E., & Dodge, K. A. (Eds.). (2005). *African American family life: Ecological and cultural diversity.* New York: Guilford.

Noguera, P. (2003). *City schools and the American Dream: Reclaiming the promise of public education.* New York: Teachers College.

Quijada, P. D., & Alvarez, L. (2006). Cultivando semillas educacionales: Understanding K–8 Latina/o students. In J. Castellanos, A. M. Gloria, & M. Kamimura (Eds.), *The Latina/o pathway to the PhD: Abriendo caminos.* Sterling, VA: Stylus.

Rogoff, B. (2003). *The cultural nature of human development.* New York: Oxford University Press.

Stolzenberg, N. M. (2001). What we talk about when we talk about culture. *American Anthropologist, 103*, 432–446.

CURRICULUM DEVELOPMENT

Curriculum development refers to the planning and documenting of a teaching/learning episode, whether it is a short course undertaken at one particular time or a long course of several years, such as a school curriculum. A school syllabus is an obvious example of a curriculum, but a 2-hour presentation or 1-hour dance class also involves consideration of what the learners should learn, what subject matter is suitable, and how that content should be communicated. Curriculum may also refer to all of the learning experiences planned by an institution or system, provided they are formally planned. The planning process is informed by educational psychology by the established information it provides developers about the nature of the learner, the nature of learning, and the nature of teaching. It also offers a challenge for educational psychology to extend existing frontiers in these areas.

This entry traces the technical process of curriculum development from a consideration of the contributing disciplines; the situational analysis or appraisal of the context; the development of the four substantive elements of objectives, content, methods, and evaluation; and the sequence followed among those elements in the process of development. It also considers the naturalistic process, or the complex interactions and negotiations between participants or developers as they search for consensus about a suitable teaching/learning episode.

The Meaning of Curriculum

One beguiling interpretation of the curriculum is to regard it as a personal transaction between a teacher/trainer and learner. Although that is true in one sense, there is also a broader context in which these transactions occur, a context that is part of the economic, political, and social nature of society. So, to understand the curriculum and its development, there is a need to understand the complex nature of society. The curriculum is not separate from society, but embedded in it.

Some definitions of curriculum often implicitly acknowledge this relationship. School curriculum, for instance, is sometimes portrayed as the context in which generations seek to establish self-identity. It is also represented as a reflection of what people in a society think, feel, and do. Other definitions that focus on a system or institution, typically the school, may equate curriculum with a syllabus prescription: an organized set of formal educational experiences, often specified by outcomes to be achieved. Others adopt a broader

interpretation, defining it as all of the learning experiences planned and guided by the school. The definition of curriculum adopted in this article is applicable to both the narrow and broad conceptions.

Decisions informing the process of curriculum design include the contributing disciplines of psychology, philosophy, and sociology; the external and internal factors comprising the situation or context; and the substantive subject matter of the curriculum itself.

The Contributing Disciplines

Knowledge of the contributing or foundation disciplines of psychology, philosophy, and sociology are essential for curriculum development. Of course, the effective developer also needs a strong understanding of people (how they grow, learn, and cope); and society (how learners are influenced by the various groups and cultures that comprise it).

Psychology contributes through its investigation of the thoughts and actions involved in teaching and learning. It not only provides information that assists curriculum development, but also contributes its own methods of investigation. The list of areas in which psychology informs curriculum planning is virtually limitless, and includes the following:

- The nature of learners—how learners react in learning situations.
- The process of thought—how learners select and process information.
- The theories of learning—how learners learn.
- The selection of learning experiences—why particular strategies are suitable for learning. Such decisions necessitate prior consideration of theories of learning, theories of cognitive and affective development, individual differences, motivation, personality, group dynamics, and teaching style.
- The conditions for learning—what constitutes optimal learning conditions (including physical requirements, grouping, and the degree of responsibility given to learners).
- The individual differences among learners—how the differences in learner ability and aptitude affect learning.
- Teacher effectiveness—how the teacher or trainer's teaching style, leadership, interaction with learners, and general behavior influence learning.
- Personality—how different learner "personalities" respond to different learning experiences.

Several psychological theorists have had an enduring impact on curriculum design. Arguably the most influential is Jean Piaget and his four-stage theory (sensori-motor, pre-operational, concrete operations, and formal operations) accounting for intellectual development from birth to adolescence. Other major contributions include the work of David Ausubel, with his general view of learning and classroom practice, and his specific notion of "advanced organizers," or strategies of organization to enhance learning; Benjamin Bloom, with his taxonomy of educational objectives in the cognitive domain; Jerome Bruner, with his models of learning; Robert Gagne, with his hierarchies of capabilities; and Abraham Maslow, with his pyramid of needs fulfillment.

The influence of Burrhus Skinner, with his theory of reinforcement and the resultant shift in emphasis to behaviorism, has been replaced more recently by a critical perspective. The current focus is on the development of critical thinking skills in learners, and the practice of thinking about thinking, and about one's own learning (metacognition). Constructivism is the new learning dynamic. Virtually at the far end of the spectrum from behaviorism, it is predicated on the basis that individual learners construct their own knowledge, rather than having it transmitted by a teacher or trainer.

The constructivist view of educational psychology is producing a revolution in the discipline, resulting in classroom and context-based research, and the development of innovative learning environments.

Sociology has had a strong functionalist influence on the development of curriculum, that is, a view of learners as the products of society, as creations of the process of socialization. This notion, sometimes viewed as an equivalent of the behavioral trend in psychology, prompted investigations of how class, gender, and ethnicity have an impact on curriculum. More recent approaches to sociology are enabling curriculum developers to detect economic, political, and social factors that create preferences for different curriculum practices. Sociology also sheds further light on future trends in society.

Philosophy contributes through the data it provides on epistemology (the nature of knowledge); ethics (the value of knowledge); and the provision of clarity (definitions, priorities, relationships between entities). It underlines the importance of developers and teachers in reflecting upon their own implicit philosophies of education to consider the consequences of this for the practice of curriculum design.

Situational Analysis

Although curriculum development is informed by the contributing disciplines, it does not take place in a vacuum. It is context bound. As curriculum is designed to "fit" the needs of particular systems or institutions, the necessary first step in the process is a situational analysis, sometimes referred to as an "environmental scam," audit, or needs assessment. This examination of the context in which the curriculum is to operate is a systematic process that continues throughout development and implementation to shape the nature of the learning experiences.

The context factors to be considered may be external to the institution or internal. These factors may include the following:

- The people for whom the curriculum is being designed, for example, the experience, cultural background, social skills, and needs of the learners, or in the case of schools, the abilities and the intellectual, emotional, social, and physical development of the learners.
- The people who are implementing the curriculum, for example the interests, expertise, strengths and weaknesses of trainers and/or the teaching style of teachers.
- The resources available to support the curriculum, for example, the equipment, facilities, space, technology, and personnel both within and external to the system or institution.
- The system or institutional requirements, for example, the mandates, system and institutional policies, and existing curricula.
- The nature of professional curricular support, for example, the availability of professional networks, consultant support, and on-the-job professional development.
- The climate, ethos, or atmosphere of the system or institution, for example, the palpable warmth, openness, and harmony, or their opposites.
- The broader societal changes and expectations, for example, the nature of technological, economic, or demographic change, or in the case of schools, the impact of multiculturalism, parent involvement, and changes in family structures.

Curriculum Elements

Whether a curriculum is developed for a day, a short course, or several years, its development requires information regarding four questions that were first posited by Ralph Tyler: What am I trying to achieve? What content will I use to achieve it? How will I organize the content to achieve it? How will I know if I've been effective?

Each of these questions involves a curriculum element or substantive element, or what have been dubbed collectively as the *curriculum commonplaces.* They are objectives/outcomes, content, methods and assessment/evaluation, and when designing curricula, developers are necessarily involved in making decisions about each of them. Following is a brief description of the four curriculum elements.

Objectives/Outcomes

There is a plethora of terms used to describe different degrees of precision in statements of curriculum intent, and the precision denoted is variable across country and context: goal, aim, general aim, specific aim, objective, behavioral objective, expressive objective, process objective, terminal objective, outcome, indicator, and pointer. The question of how precise a statement of intent should be has been a contentious issue from the 1960s, and received impetus with the advent of outcomes-based education in the 1980s and 1990s.

The proponents of using precise outcomes (and outcomes-based education) as a curriculum focus argue that it provides a more explicit indication of what students have to achieve; it diminishes the emphasis on content to be covered and accentuates the skills to be achieved; it enables the assessment of higher-order and more varied outcomes; it eliminates permanent failure (as those who don't initially achieve an outcome may do so in time); and it increases system, institution, and teacher accountability. Conversely, the detractors argue that outcomes-based education is narrow, mechanistic, and behaviorist; it constrains inquiry and creativity; it devalues the affective dimension; and it trivializes knowledge with the assumption that all learning can be broken down into outcomes.

The time-honored framework still used in schools as a guide for curriculum planning was created by Benjamin Bloom, who developed taxonomies of educational objectives in the cognitive and affective domains, respectively.

Content

The detailing of curriculum content may simply involve the identification of relevant topics, although

a curriculum of significant duration may involve the initial development of a conceptual framework, that is, a "cognitive map" displaying the relationships among the content parts (facts, ideas, skills). Frameworks may comprise subtyping, aggregation, and the relationship between entities, and are a first step in the full detailing of content. In curricula involving the development of skills, an equivalent first step is a task analysis (a breaking down of skills into more finite parts).

Methods

The methods or teaching/learning strategies are the means of acting upon the content and might include explanation, narration, demonstration, questioning, brainstorming, problem solving, role-plays, moral dilemmas, graphic organizers, values clarification, and simulations. They might also involve the structures for organizing students (working independently or using snowball or jigsaw groups). Appropriate methods are generally selected according to learning theory and other established educational principles, for example, the need for learning to be active; enable reinforcement and transfer; consider the culture and values of the learner; involve real-world tasks, informal choices, and even risk-taking; and promote the refinement of initial attempts. Benjamin Bloom's taxonomy is also used as a strategy to impart content in schools.

Assessment/Evaluation

Assessment answers Tyler's fourth question: How will I know if I've been effective? In outcomes-based education regimes, planning for assessment involves determining the extent to which outcomes have been achieved. Once regarded as occurring at the end of a course, assessment is increasingly viewed as a continuous and integral part of teaching and learning. It is also used to provide regular feedback to learners and to indicate their progress toward outcomes. The information obtained is used to inform the nature of teaching.

The debate on assessment strategies in schools remains contentious and involves the relative merits of testing (multiple-choice, true-false, short answer, matching, cloze, interpretive tests, and essays) and performance assessment (checklists, rating scales, anecdotal records, and portfolios). Some commonly accepted principles include the need for assessment to identify strengths and weaknesses, use a range of strategies, refer to explicit criteria, and provide more than one opportunity for learners to meet requirements. Arguably more contentious principles include opportunities for learner self-assessment, negotiation of required tasks, and collaboration among learners.

The mere identification of the four substantive elements does not provide information on the process of curriculum development. There are two broad types of models for developing curricula: the technical or systematic models based on a sequence (fixed or flexible) of the substantive elements of objectives, content, methods, and evaluation; and the naturalistic models based on the interactions and negotiations surrounding questions of who makes curriculum decisions and how.

Technical Models

The most common model in the field of curriculum development is Tyler's objectives model, sometimes also called the means-end model. The four questions previously posed by Tyler, from which the substantive elements are derived, are seen to provide a logical sequence of development, viz objectives, content, method, and evaluation. The first step in the model is the stating of objectives. The source of these objectives includes what students need to know and what society believes should be taught. No single source is adequate, although Ralph Tyler refers to psychological and philosophical principles as "screens" to determine the more important objectives.

The second and third steps involve the selection and organization of learning experiences. These are the means of achieving the ends or objectives. The final step of evaluation involves determining to what extent the objectives are achieved through the selected content and method.

The model has its supporters and detractors. The supporters argue three generic benefits:

1. By providing objectives as the essential first step, the model provides a clear direction for the remainder of the process. Advocates have often presented the metaphor of a journey: How often do you get into your car without knowing where you are going? An objective is equivalent to a destination.

2. The model provides a blueprint to simplify the process. The fixed sequence is regarded as a helpful structure, particularly for those who are new to the task of curriculum design.

3. The model does possess at least a prima facie logic, that is, the specification of the ends to be achieved (the objectives), followed by the means of achieving those ends (content and method).

Common criticisms of the model include the following:

- Curriculum development in practice is rarely a fixed or linear sequence. It is more typically a chaotic and idiosyncratic process with movement to and fro among the elements. Such a criticism may well be countered by the claim that typical practice is not necessarily ideal practice.
- The model does not adequately explain from where the objectives come.
- The structure of knowledge is such that it cannot be expressed in terms of prespecified performance. Theorists including Lawrence Stenhouse claim that it is possible to select content irrespective of its instrumental value in achieving ends or objectives. He argues that each form of knowledge possesses its own structure and therefore requires a distinctive methodological approach to acquiring it. This process approach acknowledges the inherent worth of certain content and accents those procedures that are appropriate for discerning the knowledge implicit in that content.
- The model risks an overemphasis on measurable objectives. Many objectives, particularly those in the affective domain, are less amenable to quantitative assessment and therefore could be ignored.

Although the objectives model is the traditional model, it is not merely a relic. The relatively recent push for outcomes-based education may well have rejuvenated an objectives model in which outcomes are substituted for objectives.

The interaction or dynamic model indicates a different relationship among the curriculum elements. Curriculum development is regarded as a dynamic process that can begin with any curriculum element and follow any sequence among them. Instead of the elements following a fixed sequence, they are regarded as interactive and progressively modifiable (a change to one element will necessitate changes to the other elements). For example, the addition of content may involve changes to objectives, methods, and evaluation.

The claimed strengths and weaknesses are not as well documented as those of the objectives model, but supporters cite the strengths in terms of the characteristics—the model is psychologically efficient and flexible in enabling the developer to begin with any one of the elements, follow any sequence, and move to and fro among them. The lack of being constrained by a fixed process is regarded as a truer reflection of the reality of curriculum development.

Proponents of the objectives model believe that not specifying objectives as the starting point may result in their being accommodated to the content specification and thereby be contrived and decorative. They also argue that if the model is not systematic, that is, has no fixed sequence, the constant changes of direction might be reflected in a piecemeal finished product or curriculum.

Naturalistic Models

Only relatively recently has the curriculum development literature moved beyond a consideration of the substantive elements to an examination of the complex interactions, negotiations, and compromises surrounding questions of who makes curriculum decisions and how. One such early model is that of Decker Walker, who coined the term *naturalistic* to portray how curriculum planning actually occurs. His three-step sequence of platform, deliberation, and design is a reflection of observed practice.

The term *platform* refers to the shared principles that evolve from the discussion of participants about their various values, beliefs, perceptions, and commitments. For Walker, this suggests both a political platform and something to stand on, and includes an idea of what is and a vision of what ought to be. A platform might comprise conceptions (beliefs about what exists), theories (beliefs about relations between existing entities), aims (beliefs about what is desirable), images (less well-formulated notions that a change is desirable yet with no clear indication of what), and procedures (vague indications of proposed action with no clear specification of why they are desirable).

Deliberation involves the ways in which beliefs may be used to identify problems with the existing curriculum, and how new curricula may eliminate these problems. So, deliberation comprises interaction and argument in debating alternative perceptions and solutions. It may involve the expression of feeling and frustration as developers express their different ideas of what should be done. Walker describes the process as chaotic and confused.

Design begins once deliberation has made explicit what has to be done, that is, when the participants have achieved consensus about beliefs and viable solutions.

The model accents the need for meaningful debate and the sharing of beliefs, feelings, and perceptions. It also allows for the confrontation of conflicts (the stimulus for change). One prima facie criticism is that the model may be more applicable to large-scale curriculum design involving many participants, and for which deliberation is a prerequisite, as opposed to curriculum change for which consensus is assumed and that involves routine tasks.

There is possibly also a danger in viewing curriculum development as the reconciling of conflicting and competing interests, rather than promoting the perception of a core of common interests. Such a view does not diminish the importance of difference or conflict. The latter is regarded as a prime factor in promoting quality curriculum development as it causes decisions to be determined more thoughtfully; stimulates creativity by exposing different views or solutions; increases psychological intensity, physical energy, and curiosity; produces awareness of problem areas in relationships; and enhances self-knowledge. Democratic societies celebrate difference and productive conflict, but a key to understanding the curriculum is to discern common interests that unite people.

Stakeholders

Achieving consensus may be more difficult if the stakeholders, or those who have an interest in the likely impact of the curriculum, come from a broad spectrum of society. Arguably, the values and views of participants may be more homogeneous in a single, small institution in which the services are directed toward a well-defined goal. Conversely, there are individual and group interests in school curricula that are both internal and external:

- Parents have personal, social, and often vocational aspirations for their children, and they invest their faith in the school and its curriculum to ensure that this happens.
- Teachers are disposed toward the academic rather than the vocational, and possibly the theoretical rather than the practical. Their task is to interpret system guidelines and to add a pedagogical dimension that creates ongoing curriculum experiences for students.
- Community groups, notably social services agencies dealing with social, medical, and welfare issues related to young people and their families, have a stake in school curriculum. As students have lives outside classrooms, their external environment is a powerful influence on attitudes and behavior, and the school curriculum is a potentially powerful force in developing appropriate behaviors. In most countries, groups of individuals have special interests that school curricula must also address. Such interests may include indigenous groups, people with disabilities, people from a non-English-speaking background, people who live in poverty, and the geographically isolated.
- Governments have interests that are largely, although not exclusively, economic. The nature of the school curriculum will determine the knowledge and skills that future citizens will possess and hence their capacity to contribute to the nation's economy. Of course, democratic governments also desire a community that is socially cohesive, politically literate, culturally sophisticated, and just.
- Businesses share the government's economic interest. They need workers who are both skilled in particular areas and also literate and numerate.
- Universities and agencies of further education have a strong interest in the nature of the school curriculum, and often play a watchdog role to ensure that potential students are well equipped to undertake further study.

If the school curriculum is perceived as a means by which students become responsible and productive citizens, then virtually everyone has a stake in the nature and development of the curriculum.

Laurie Brady

See also Bloom's Taxonomy of Educational Objectives; Instructional Objectives; Learning Objectives; Teaching Strategies

Further Readings

Ausubel, D. (1960). The use of advanced organizers in the learning and retention of meaningful verbal learning. *Journal of Educational Psychology*, *51*, 267–272.

Bloom, B. S. (1956). *Taxonomy of educational objectives: The classification of educational goals: Handbook 1, cognitive domain.* New York: David Mackay.

Brady, L., & Kennedy, K. (2007). *Curriculum construction* (3rd ed.). Sydney: Pearson Education Australia.

Bruner, J. (1963). *The process of education.* Toronto: Random House.

Gagne, R. (1965). *The conditions of learning.* New York: Holt, Rinehart and Winston.

Maslow, A. (1968). *Toward a psychology of being.* Princeton, NJ: Princeton University Press.

Piaget, J. (1959). *The language and thought of the child* (M. Gabain & R. Gabain, Trans.). London: Routledge and Kegan Paul.

Smith, D. L., & Lovat, T. J. (2003). *Curriculum: Action on reflection* (4th ed.). Tuggerah, NSW: Social Science Press.

Stenhouse, L. (1975). *An introduction to curriculum research and development.* London: Heinemann.

Tyler, R. W. (1949). *Basic principles of curriculum and instruction.* Chicago: University of Chicago Press.

Walker, D. F. (1971). A naturalistic model for curriculum development. *School Review*, *80*(1), 57–65.

D

The price of the democratic way of life is a growing appreciation of people's differences, not merely as tolerable, but as the essence of a rich and rewarding human experience.

—Jerome Nathanson

Deductive Reasoning

Psychological investigations of deductive reasoning focus on inferences drawn from premises conveying operators such as *if*, *and*, *or*, and *not*, and quantifier terms such as *all* and *some*. Formal logic typically is used to determine which deductive inferences are valid. People routinely draw many valid inferences, but often accept others that are fallacious, and psychological research needs to account for both. Some researchers, such as Jonathan Evans, have focused on task features that influence fallacious inferences, proposing biases and heuristics in reasoning. This approach, however, does not explain the valid inferences that people make, nor the intuitions about the validity of such inferences.

Researchers such as Patricia Cheng, Keith Holyoak, and Leda Cosmides have focused on the role of content in making inferences, and certain types of content can both discourage some common fallacies and encourage some valid inferences. This approach, however, does not explain a wide set of valid inferences made with content outside of those described in these theories. Other researchers, such as David Over, Mike Oaksford, and Nick Chater, have proposed Bayesian inferential processes, in which knowledge about problem content

guides computations about probabilities. These theories do not explain inferences that are made from premises with abstract content.

Two research paradigms address deductive inferences across all sorts of content: mental-logic theories and mental-models theories. Two mental-logic theories have been proposed, one by Lance Rips and another by Martin Braine and David O'Brien. These theories propose sets of inference schemas and procedures for implementing the schemas. Many reasoning errors are explained by limitations in the reasoning procedures. Braine and O'Brien proposed a universally available direct-reasoning routine that applies the inference schemas whenever the requisite propositions are considered simultaneously. For example, when premises of the form *p or q* and *not p* are in working memory conjointly, a schema infers *q*, and when premises of the form *if p then q* and *p* are considered conjointly, another schema infers *q*. Considerable evidence shows that people apply such schemas automatically and effortlessly whenever the premise information is available both in laboratory problems and in stories. Reasoning strategies that go beyond the basic program are not universal and require effort, and problems that require such strategies are solved far less and often lead to fallacious inferences. Another source of reasoning error addressed by the mental-logic approach is found in pragmatic inferences. For

example, *if p then q* can invite one to infer *if not p then not q*, which is not deductively valid.

Philip Johnson-Laird and Ruth Byrne have proposed a mental-models theory in which people construct analogical representations of premise information. Because working memory has limited capacity, people tend to construct only minimal representations, although a completely logical representation usually would require more complete models. For example, given the premises *if p then q* and *p*, one can construct the following incomplete models:

$$p \quad q; \quad p; \quad p \quad q$$

where the models for the two premises and then for their combination are separated by semicolons. A conclusion will be based on the final model, which would support a conclusion of *p and q*, although no one states this conclusion, saying instead only *q*. Models theorists explain this by arguing that because *p* is a categorical premise, its inclusion in a conclusion would violate pragmatic constraints.

Models theory explains reasoning errors principally in terms of the number of models required to support a conclusion. For example, given *if p then q* and *q* as premises, one is likely to form these models:

$$p\,q; \quad q; \quad p \quad q$$

which support the fallacious inference of *p* that people often endorse. The inference could be withheld, however, if people "fleshed out" the models to include all possibilities, but this would tax working memory, so people tend to endorse the fallacious inference based on the incomplete models.

Mental-models theory has an advantage in that it promises to provide a general theory of reasoning, that is, being applicable not only to deductive reasoning, but also to all areas of reasoning. The mental-logic approach makes no such sweeping claim and is limited to providing an account of how logical inferences are made for propositional operators such as conjunctions, alternatives, negations, and suppositions, and for various quantifiers. Mental-logic theory, however, has the advantage of making specific predictions that are supported by the data, whereas mental-models theory leads to some problematic predictions. For example, consider the premises mentioned above, *if p then q* and *p*, which in models theory would lead to the conclusion *p and q* except that *p* is not included in the stated conclusion because of pragmatics. It

follows that a similar problem with the same premises but with *not both p and q* as a conclusion should lead to a straightforward evaluation of "false" because *p* and *q* are in the final model. People, however, make an intermediate inference of *p* before saying "false," which models theory cannot explain. (The intermediate inference is explained by mental-logic theory.) Of course, the two theories are not exclusive, and one need not choose one over the other. A skilled thinker would benefit from using both sorts of procedures.

David P. O'Brien

See also Inductive Reasoning; Intelligence and Intellectual Development; Learning; Learning Style

Further Readings

Braine, M. D. S., & O'Brien, D. P. (Eds.). (1998). *Mental logic.* Mahwah, NJ: Lawrence Erlbaum.

Cheng, P. W., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology*, *17*, 391–416.

Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition*, *31*, 187–276.

Evans, J. St. B. T., & Over, D. E. (2004). *If.* New York: Oxford University Press.

Johnson-Laird, P. N., & Byrne, R. M. J. (1991). *Deduction.* Hillsdale, NJ: Lawrence Erlbaum.

Johnson-Laird, P. N., & Byrne, R. M. J. (2002). Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review*, *109*, 646–678.

O'Brien, D. P. (2004). Mental-logic theory: What it proposes, and reasons to take this proposal seriously. In J. Leighton & R. J. Sternberg (Eds.), *The nature of reasoning* (pp. 205–233). New Haven, CT: Yale University Press.

Rips, L. J. (1994). *The psychology of proof: Deductive reasoning in human thinking.* Cambridge: MIT Press.

Descriptive Statistics

Descriptive statistics constitute a branch of statistics that provides a variety of techniques used to present a quantitative summary of a given set of data. In contrast to descriptive statistics are inferential statistics, which provide ways to make inferences about the population from which the data were sampled.

When summarizing data on a single variable is of interest, there are at least two major characteristics to be considered: central tendency and dispersion. These

characteristics represent some aspects of the distribution of the data. The *central tendency* denotes where the center of the data distribution is located. The *mean*, the *median*, and the *mode* are commonly used to describe the central tendency. The mean is the average of all data values, that is, the sum of data values divided by the number of observations. The median is the midpoint in a set of sorted data values; half of the data values fall above the median and half of the data values fall below it. The mode is the data value that appears most frequently in the given set of data.

The *dispersion* refers to the degree to which the data vary around the center of the distribution, or to what extent the data are spread out. One of the measures of dispersion is the *standard deviation*, which indicates the extent to which individual data points depart from the mean on average. Technically, it is defined by the following formula:

$$s = \sqrt{\frac{\sum_{i=1}^{N} (X_i - \bar{X})^2}{N}},$$

where X_i is the ith observation, $\bar{X}$ is the mean, and N is the total number of observations.

Another term for dispersion is variation. A measure for data variation is variance, which is defined as the square of the standard deviation. To illustrate how these indices are used, a set of data will be analyzed.

Example 1

Suppose that 20 children took a math test and we observed the following scores:

$$6, 12, 11, 14, 11, 13, 14, 11, 11, 7, 8, 10,$$
$$12, 15, 7, 8, 12, 16, 5, 11$$

The mean for the above data is 10.70. If the above data are sorted, scores for the 10th and 11th observations are both 11. The median is their average (because the total number of observations is an even number), and thus calculated as 11. The mode for the above data is 11, because the score 11 is observed most frequently (five times). When the data distribution has a single mode and is almost symmetric about its center, as in this example, these three measures give almost the same values. When the distribution is skewed or when there are a few extreme values (outliers) in the data, the three measures of central tendency can disagree.

Although the mean is most commonly used because of its computational and interpretational convenience, the median or the mode is preferred for skewed data or when outliers are present, because the mean is sensitive to these factors and could be misleading as a representative value of the distribution. The standard deviation is calculated as 3.05 for the above data and the variance is thus 9.30. It should be noted that the standard deviation, like the mean, may not be representative of data if the distribution is highly skewed (i.e., lacks symmetry), because the average dispersion can differ for the data points above the mean and for those below the mean (and the mean itself is also affected by the skewness).

Quantiles such as *percentiles* and *quartiles* are also useful in describing a data distribution. Let the data be sorted in an ascending order. Divide the sorted data into 100 subsets of equal size. The kth percentile is defined as the boundary value between the kth subset and $(k+1)$th subset. The interpretation of the kth percentile is that $k\%$ of the observations in the data fall below the kth percentile. The zero percentile corresponds to the smallest value in the data, and the 100th percentile to the largest value. Quartiles are defined in the same manner, but the data are divided into four subsets. Thus, the first, second, and third quartiles correspond to the 25th, 50th, and 75th percentiles, respectively. The median is equivalent to the 50th percentile or the second quartile.

For the data in Example 1, the first, second, and third quartiles are 8, 11, and 12.25, respectively, indicating that half of the scores are between 8 and 12.25 and that the data distribution is somewhat skewed to the left, because the distance between the first and second quartiles is longer than the distance between the second and the third. The *interquartile range* is defined as the difference between the third and the first quartiles, and it is used as a measure of dispersion. In the above case, the interquartile range is 4.25.

When two or more variables are summarized at the same time, one may be interested in describing relationships between those variables as well as univariate summaries of individual variables. The *correlation* is one of the most commonly used measures of a relationship between two variables. The correlation between two variables X and Y is defined by the following formula:

$$r = \frac{\sum_{i=1}^{N} (X_i - \bar{X})(Y_i - \bar{Y})/N}{\sqrt{\sum_{i=1}^{N} (X_i - \bar{X})^2/N} \sqrt{\sum_{i=1}^{N} (Y_i - \bar{Y})^2/N}}.$$

The numerator in the above formula is called the *covariance*, and the denominator is the product of standard deviations. The correlation measures the strength of a linear relationship between two variables. A correlation of 1.0 (or –1.0) indicates the perfect positive (or negative) linear relationship (i.e., all data points lie on a single line in a two-dimensional plot). A zero correlation implies that there is no linear relationship between two variables.

Example 2

Suppose that you have scores of a reading test from the same 20 children as in Example 1:

6, 14, 14, 9, 7, 18, 10, 9, 10, 9, 10, 9, 15, 15,

10, 9, 13, 12, 11, 12

The standard deviation of these scores is 2.99. The covariance between math and reading test scores is 4.40, resulting in the correlation .48. Thus, reading and math test scores are positively correlated.

Graphical and tabular representations are other ways to present summaries of data. For example, a data distribution of a single variable can be visualized by a *histogram.* A *frequency table* may be often used to summarize categorical data. A relationship between two variables is made visible by drawing a *scatterplot.* These representational tools complement the various descriptive statistics that may serve as aides to understand data in more detail.

Kentaro Kato and William M. Bart

See also Inferential Statistics; Normal Curve; Standard Scores

Further Reading

Glass, G., & Hopkins, K. (1996). *Statistical methods in education and psychology* (3rd ed.). Boston: Allyn & Bacon.

DIAGNOSTIC AND STATISTICAL MANUAL OF MENTAL DISORDERS

The *Diagnostic and Statistical Manual of Mental Disorders* (*DSM*) is a comprehensive summary of all known forms of psychopathology. It is published by the American Psychiatric Association and is currently in its fourth edition (*DSM–IV*). The *DSM* is used by both researchers and clinicians to identify psychological disorders.

History

The need for statistical information was the initial impetus in the United States to develop a classification system of mental disorders. The first official attempt was the recording of the frequency of one category labeled "idiocy/insanity" in the 1840 census. By the 1880 census, seven categories of mental illness were distinguished—dementia, dipsomania, epilepsy, mania, melancholia, monomania, and paresis. In 1952, the American Psychiatric Association Committee on Nomenclature and Statistics published the first edition of the *DSM*. This edition and the one that followed were based on a psychodynamic interpretation of each diagnosis. Beginning with the third edition, the *DSM* has used an atheoretical approach to diagnosis, which has increased the reliability between raters for establishing diagnoses.

Use of the *DSM*

The *DSM* consists of a multiaxial classification system used by clinicians to diagnose individuals, plan treatment interventions, and predict probable outcomes. Diagnoses are established along the following five axes:

Axis I: Clinical Disorders and Other Conditions That May Be a Focus of Clinical Attention

Axis II: Personality Disorders and Mental Retardation

Axis III: General Medical Conditions

Axis IV: Psychosocial and Environmental Problems (rated with descriptive categories)

Axis V: Global Assessment of Functioning (rated from 1 [persistent danger to self or others] to 100 [superior functioning to a wide range of activities])

All mental disorders/conditions listed in the *DSM* fall under Axis I with the exception of Personality Disorders and Mental Retardation (listed under Axis II). Mental disorders or conditions that are the main focus of clinical attention are usually listed on Axis I. This axis is composed of the following categories: schizophrenia and other psychotic disorders, mood disorders,

anxiety disorders, somatoform disorders, disorders usually first diagnosed in childhood (e.g., autistic disorder), eating disorders, delirium, dementia, amnestic and other cognitive disorders, sleep disorders, impulse control disorders, and adjustment disorders.

Axis II is used for reporting personality disorders as well as mental retardation. Listing these conditions on a separate axis emphasizes the chronic and inflexible nature of this class of disorders. The personality disorders are classified into three broad clusters, as follows: cluster A, odd and eccentric (paranoid, schizoid, and schizotypal); cluster B, expressive/labile (antisocial, borderline, histrionic, and narcissistic); and cluster C, anxious (avoidant, dependent, and obsessive-compulsive).

General medical conditions that may be pertinent to the understanding and/or treatment of an individual's mental condition are reported on Axis III. For example, persons with severe forms of arthritis are more likely to be depressed due to the pain and physical limitations associated with their medical condition.

Axis IV accounts for environmental or psychosocial factors that may potentially limit the effectiveness of the treatment for disorders on the first three axes. For example, the loss of job and the financial issues associated with that loss would complicate treatment of depression. There is a broad range of possibilities that could be listed under this category. The clinician should list the category as well as specifics about the problem. The main categories with some examples are as follows: problems with primary support group (e.g., divorce, neglect of child, health problems among family members); problems related to the social environment (e.g., death of a friend, discrimination, adjustment to retirement); educational problems (illiteracy, academic problems); occupational problems (unemployment, stressful work environment, discord with a co-worker or boss); housing problems (homelessness, unsafe neighborhood); economic problems (poverty, inadequate finances); problems with access to health care (lack of transportation to health care services, inadequate or lack of health insurance); problems related to interaction with the legal system/crime (incarceration, litigation, victim of crime); other psychosocial and environmental problems (exposure to disaster, discord with nonfamily caregivers such as counselor, social worker, or physician).

Axis V is used by the clinician to give an overall picture of the client or patient's level of functioning in recent and current everyday living. Thus, this axis is rated for the highest level estimated for the year prior to the assessment as well as the functioning level at the time of the diagnosis. The rating for the highest level of functioning in the previous year often helps the clinician provide a prognosis. The scale to make the ratings is the Global Assessment of Functioning Scale (GAF), which is divided into 10 ranges of functioning from 0–100 (0–10; 11–20; 21–30; 31–40; 41–50; etc.). The clinician considers psychological, occupational, and social functioning to assign the patient a particular number within the 0–100 range that coincides with a description with that particular number range. For example, a patient's GAF score of 70 would represent "mild symptoms (e.g., depressed mood or mild insomnia) *or* some difficulty in social, occupational, or school functioning (e.g., occasional truancy, or theft within the household) but generally functioning pretty well, has some meaningful interpersonal relationships" (American Psychiatric Association, p. 34).

Importance

The overall purpose of the multiaxial format of the *DSM* is to provide a comprehensive and organized system for categorizing various types of psychopathology. This system allows different professionals (e.g., psychologists, physicians, social workers) to communicate effectively regarding a patient's mental disorder and associated conditions. It also provides information to third-party payers.

Karen D. Multon and Daniel Suitor

See also Communication Disorders; Conduct Disorders; Disabilities; Dyslexia; Eating Disorders; Intelligence Quotient (IQ); Learning Disabilities; Mental Retardation

Further Readings

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., Text rev.). Washington, DC: Author.

Meyer, R. G. (2006). *The clinician's handbook: Integrated diagnostics, assessment, and intervention in adult and adolescent psychopathology* (4th ed.). Boston: Allyn & Bacon.

Reid, W. H., & Wise, M. G. (1995). *DSM-IV training guide.* New York: Brunner/Mazel.

DIRECT INSTRUCTION

Direct instruction has at least two meanings. The first meaning refers to the generic behavioral features of explicit teaching in which teachers demonstrate a performance, guide students as they attempt to perform, and continue student/teacher exchanges until students can perform without further guidance. Barak Rosenshine and others have described these features in detail. Thomas Gilbert developed a classic direct instruction model, which he called *Mathetics*, a Greek word meaning "to learn." Teachers who implement direct instruction methods are usually teaching well-defined learning objectives; for example, simple associations such as the table of elements in chemistry; sequences such as the long division algorithm; concepts such as kinds of trees or modern art styles; and applications of principles such as supply and demand in economics, rules of punctuation, detecting an author's purpose, or inferring more information about characters and settings in literature. Direct Instruction teachers are usually aiming for student mastery of objectives. This entry refers to these generic features of instructional delivery as *direct instruction*, with lower-case letters to begin each word.

A second meaning of *Direct Instruction* refers to a theory of instruction developed by Siegfried Engelmann, specific programs designed from his theory, and specific direct instruction delivery techniques espoused by his theory and programs. The first letter in each word, *Direct Instruction*, is capitalized when referring to Engelmann's theory, programs, and methods. Direct Instruction is often abbreviated as DI, with the two letter-names pronounced when referring to it.

First, this entry describes features of generic direct instruction presentation methods, and then describes Engelmann's Direct Instruction.

Presentation Methods

A direct instruction lesson is an exchange between a teacher and either a single learner, a small group of learners, or an entire class. Students engage with a teacher in highly interactive lessons that introduce one performance or skill at a time and cumulatively combine them as accuracy emerges. Establishing new performances with direct instruction is equivalent to generalized imitation training. Teachers *demonstrate* and *model* expected performance, *lead* and *prompt* students through the performance, and then *release* or *test* the students' performance. These steps in the process are easily identified during instruction when teachers preface by saying "My turn" when they demonstrate and model, "Our turn" when they help the learners to perform accurately, and "Now your turn" when they check to see whether students can perform without assistance.

Many learners need a very explicit and structured version of direct instruction to make rapid progress. They also need a curriculum that has been more carefully designed than the average textbook presents or the average teacher outlines in a lesson plan. These include (a) young learners, (b) older learners with deficient learning skills, (c) learners with learning and developmental disabilities, (d) learners at a cultural disadvantage in academic learning, and (e) learners who encounter brand-new content that allows for little transfer of training. For these learners, Engelmann's Direct Instruction (DI) curriculum, as well as his direct instruction procedures for delivering instruction, are very effective. His direct instruction procedures are examined first, and then his Direct Instruction (DI) program design.

Teachers present to learners DI lessons that designers have scripted, word for word, in advance. The scripts take the utmost care to present a hierarchy of component and cumulative compound skills. The best DI scripts have been "learner verified"; that is, tested with hundreds of learners in many settings and revised until they work. During a DI lesson, learners answer teacher-initiated questions and tasks in unison. Teachers move through a series of questions and tasks in a predictable manner, provide attending cues to ensure that all learners are focused on the task, and signal learners to ensure that they respond together. Teachers may use a variety of signals—for example, tapping a pen on an overhead transparency or snapping their fingers—until they find one that works, is comfortable to use, and doesn't distract students. Teachers systematically vary the latency between the question and the response signal to ensure that the signal tightly controls responding.

Teacher and students volley many times a minute with their questions and answers. Teachers praise and correct student responses, providing parallel tasks until all students are accurate. The explicitness and careful progression of Direct Instruction lessons ensures that at-risk students develop flawless skills very quickly.

With DI lessons, not only do low-skilled learners master the skills and concepts, but also inexperienced teachers learn to establish new performances. In the process of using DI scripts—which tell the teacher exactly what to say, how the learner should respond, how to predict errors, and how to provide error correction procedures—teachers are explicitly shaped to establish skills.

The challenge for the inexperienced Direct Instruction teacher is to faithfully present the lesson as designed. This requires achieving choral responding among learners, listening carefully to the quality of their responses, providing encouraging feedback following correct responses, and applying specific error correction procedures that effectively reduce errors. They must proceed at a pace that provides optimal thinking time before students respond. The pace must be gradually increased to squeeze out most thinking time as skills reach mastery. Teachers must keep track of the three delivery phases (demonstrate, prompt, and test) and engage in them recursively. For example, teachers may demonstrate a skill a few times, then try an "our turn" exchange. If it flops, then the teacher must return to the demonstration phase. Likewise, "your turn" attempts may fail, at which point the teacher must return to "our turns." Teachers must also provide individual turns to students, to make sure that they did not miss errors during choral responding. Even though teachers may consult the script during the lesson, they must be conversant enough with its pattern to maintain eye and ear contact with learners. The best results occur when teachers are able to implement programs with a great deal of procedural reliability, although the lessons are generally forgiving about small lapses in procedure.

Direct Instruction Program Design

Engelmann describes seven criteria for Direct Instruction program design. First, average to below-average teachers should be able to teach a program successfully after receiving a week or less of training with some follow-up, in-classroom coaching. Second, teachers, principals, and other instructional leaders should be able to reliably predict the amount of student progress a program will produce in a given time period. At least three features of Direct Instruction programs allow for such reliable predictions. Because students are homogeneously grouped for instruction, the teacher can focus on teaching the same content to all students. Homogeneous grouping makes the instructional agenda very clear and has remarkable data supporting it independent of DI. Given a clear agenda, the lessons in a program should be scripted to give teachers the clearest and most consistent language for teaching the material, as determined in empirical tryout and revision of the scripted lessons, and to keep teaching focused on the content to be mastered and not extraneous but related matter. Scripts include many examples of a learning objective, as well as close-in nonexamples ("almost like but not quite") to teach the *general case* of a concept, skill, or principle. The lessons in a Direct Instruction program also have a specified time during which students are expected to master the content, typically 30 to 40 minutes in length. Research indicates that high performers typically master about 1.3 lessons per day, middle groups about 1 lesson a day, and lower groups about .7 lessons per day.

The third criterion for Direct Instruction program design is that the sum of the skills and concepts to be taught is substantially greater than the sum of the skills the students currently master. Inspection of DI programs may indicate less variety of activities or topics when compared to traditional materials, but there is usually no evidence that the traditional materials actually teach the content they purport to teach. The empirical nature of the design, tryout, and revision process of Direct Instruction programs promotes mastery.

Fourth, Engelmann's Direct Instruction programs are based on a theory of instruction, which is beyond the scope of this entry. Suffice it to say, his logical analysis of knowledge, analysis of teacher–student communication, and functional analysis of student behavior has produced many effective programs. Empirical evidence shows that the lesson presentations are consistent with one, and only one, interpretation.

The fifth criterion for Direct Instruction program design is that interactivity must be built into the lessons to guide the teacher in adjusting the rate and type of presentation students need at each moment. Because of the high frequency of teacher–student exchanges, teachers know when to engage in each of the three direct instruction delivery phases (i.e., provide more demonstration, provide prompts to guide the students, and provide student tests). A DI program should also include periodic cumulative, curriculum-based tests.

The sixth criterion is the long view that a curriculum sequence takes, anticipating the prerequisites for

later, more complex learning. A program must fully prepare the student early and systematically with skills and concepts that will guarantee the student learns later, more complex content. This focus on prerequisite skills promotes teaching priority skills over trivial skills. The seventh criterion states that a Direct Instruction program should provide consumer protection information to suggest that successful outcomes are possible. The thorough field testing that each Direct Instruction program receives virtually guarantees that teachers will teach successfully and students will learn to mastery.

Engelmann and his colleagues have published more than 50 Direct Instruction programs, all classroom-tested and revised with thousands of students until they work. Examples include *Reading Mastery, Corrective Reading, Language for Learning, Language for Thinking, Expressive Writing, Spelling Mastery, and Connecting Math Concepts*. These programs are available through SRA/McGraw-Hill.

Direct Instruction Evaluation

The U.S. Department of Education first documented the effectiveness of Direct Instruction in the 1960s. They sponsored Project Follow Through ("following through" on the preschool Head Start programs). Anyone with an educational model for teaching disadvantaged students in grades K–3 was invited to enter a comparative evaluation; 13 different educational models participated. Twenty-one community sites selected the Direct Instruction model. The Direct Instruction students outperformed students in all other Follow Through models including Open Education, Behavior Analysis, Bank Street, Cognitive Curriculum, and Parent Education. The DI students showed superior progress in reading, arithmetic, spelling, and language. Direct Instruction students also had the most positive self-image of any students.

Since Project Follow Through, 151 studies have compared Direct Instruction to another teaching method, the largest number of studies done on any method of teaching in the history of educational research. One hundred and nine, or 64% of the studies, statistically favored Direct Instruction. Fifty-nine, or 35% of the studies, showed no difference, and only two, or 1%, favored a non–Direct Instruction method. A meta-analysis of Direct Instruction studies revealed an effect size of .97, rare in educational research. It is also remarkable that these studies involved a full range of student populations—preschoolers; elementary and secondary students; regular and special education students; adults; and various minority populations, including non-English speakers. No other instructional method that we know has produced more favorable results than Direct Instruction.

Direct Instruction Revisited

The importance of tighter, DI-scripted lessons for at-risk learners of all kinds, as well as for inexperienced teachers, has been discussed. When students are average to above-average learners, and teachers have direct instruction experiences, generic direct instruction lessons to establish performances may be successful. At least half or more learners have more advanced learning-to-learn skills than their at-risk peers. Most courses of study are not brand-new experiences for a learner, although some foreign language and science courses are examples of exceptions. Teachers often reach the point in training and practice where they have fluent generalized repertoires of effective instructional approaches. Instructional lessons for these students and with these teachers allow a shift to more generic direct instruction approaches as described by Gilbert, Rosenshine, and others.

Generic direct instruction lessons still involve frequent exchanges between teacher and students, with praise and error corrections. They also contain model, prompt, and test modulations until all students can perform successfully without assistance. However, students may encounter more teacher talk before responding, with the teacher thinking aloud to show how he or she formulates a successful answer to a task. For example, to teach a reading comprehension skill "making an inference," the teacher may define *inference*, read a short passage, and "think aloud" one or more inferences that can be drawn from the passage. Learner responses to tasks may also require lengthier answers. The teacher may read another short passage and ask students to make inferences and justify their answers. Unison responding may not be choral; teachers may give "think time" to students to write answers to such tasks, and then call on a student or two to read their answers, while others check their answers against the feedback the teacher provides. Students may also compare their written answers in peer pairs. The teacher may back up and provide more think-aloud models if the students are off-base. The teacher may provide hints and otherwise lead

students to make more plausible inferences, and then read a new short passage and test the student's skill at making inferences. Many variations are possible. Mastery is still the learner's goal, and learner performance still drives the teaching, but the learner is expected to reach mastery with less explicit teacher shaping.

Kent Johnson and Elizabeth M. Street

See also Effective Teaching, Characteristics of; Explicit Teaching; Special Education; Teaching Strategies

Further Readings

Adams, G. L., & Engelmann, S. (1996). *Research on Direct Instruction: 25 years beyond DISTAR.* Eugene, OR: Association for Direct Instruction.

Engelmann, S., & Carnine, D. W. (1982). *Theory of instruction: Principles and applications.* Eugene, OR: Association for Direct Instruction.

Gilbert, T. F. (1978). *Human competence: Engineering worthy performance.* New York: Wiley.

Johnson, K. R., & Layng, T. V. J. (1992). Breaking the structuralist barrier: Literacy and numeracy with fluency. *American Psychologist*, *47*, 1475–1490.

Johnson, K., & Street, E. M. (2004). *The Morningside model of generative instruction: What it means to leave no child behind.* Concord, MA: Cambridge Center for Behavioral Studies.

Markle, S. M. (1990). *Designs for instructional designers.* Seattle, WA: Morningside Press.

DISABILITIES

In the 2003–2004 school year, the federal government reported that more than 6.6 million youth with disabilities were served by federally funded programs in public schools. These youth include individuals with learning disabilities, speech/language impairments, developmental and intellectual disabilities, emotional disturbance, hearing impairments, orthopedic impairments, other health impairments, visual impairments, deafness, blindness, autism, traumatic brain injury, developmental delay, and multiple disabilities. This entry discusses perspectives on disability and the intersection of disability and educational contexts, including pertinent legislation and issues related to the diagnosis and education of youth with disabilities.

Perspectives on Disability

Disability has been conceptualized in terms of type, function, and social construction. Most commonly, perhaps, disability has been construed as a type of limitation with particular characteristics. The four major types are: (1) physical (e.g., spinal cord injury, polio); (2) sensory (e.g., deafness, blindness); (3) cognitive (e.g., learning disability, intellectual disability); and (4) emotional (e.g., schizophrenia, bipolar disorder). A full typology would indicate causes, characteristics, courses, and outcomes for each type of disability and would help to identify the kind of disability one has. Although the medical profession and others have found this approach a useful way to organize knowledge about disability, typologies have not always been able to live up to their promise of providing such information, nor have they always been able to capture the range and kind of disabilities human beings experience.

A second approach is functional. Here the goal is to specify the actual impairments the disabling condition imposes. For example, the Americans with Disabilities Act includes a definition of disability as a condition that results in an individual being unable to perform basic life functions independently, such as walking and bathing. This approach is valuable for specifying the nature and threshold level of impairment involved with the disability. The U.S. legal system sometimes uses a functional approach to identify individuals with disabilities who qualify for protection and benefits under a particular law or policy without regard to whether the exact type of disability they may have is well defined or understood.

The socially constructed approach recognizes the role of society in defining disability. A basic principle is that a disability is defined as an individual's inability to perform in some socially valued way. Thus, if a society values intellectual work, then an individual who cannot learn to read may be considered to have some kind of a cognitive disability. However, in an agrarian society, if most of the population is not able to read, then cognitive disabilities may not be considered important. In fact, many less severe disabilities considered high incidence today, such as learning disabilities, were not recognized as disabilities prior to the Industrial Revolution and its attendant urbanization. A second basic principle of the socially constructed approach addresses the important role of societal attitudes in shaping the nature of the

disability experience. In different societies, attitudes toward disabilities and the persons who have them may be negative, neutral, and/or positive. For example, in some African cultures, an individual who is hallucinating may be considered to be divinely inspired and capable of special healing. In contrast, in mainstream U.S. culture, such a person may be considered to be in need of medication and perhaps hospitalization to treat serious mental illness. More generally, people with disabilities may report that in the eyes of others, their disability overshadows their abilities. Consequently, the attitudes of others toward people with disabilities may constitute a greater difficulty than any impairments resulting from their disability per se.

In schools and other educational contexts, all three of these approaches to defining disability are salient. The attitudes of society, as represented by the perspectives of principals, teachers, and student peers, shape the education of students with disabilities. Where these attitudes are positive, students with disabilities may be more welcomed and fully included in the academic, extracurricular, and social life of the school. In these schools, research suggests that students with disabilities are often more likely to develop positively and perform well socially, psychologically, and academically. Because of the information age in which we live, education is considered an essential opportunity in U.S. society. Therefore, disabilities that affect one's education are socially constructed to be of great importance. Consequently, there are laws protecting the educational opportunity of students with disabilities and professionals to diagnose these disabilities and support affected students and their families. School psychologists and other professionals emphasize both the type of disability and function when they diagnose students to ascertain their eligibility for special education services and determine the intervention and support likely to be most helpful for students with disabilities and their family members.

History of Educational Practices and Legislation

The educational rights of and opportunities for youth with disabilities have changed substantially over time. Early practices advocated for the removal of children with visible types of disabilities from their communities to institutions, thereby reducing the need for general educational systems to concern themselves with their education. Before educational systems began to seriously consider the intersection of disability and education, a few individuals became interested in the topic (e.g., Itard in France and Montessori in Italy). As early as the 1800s, schools for children with sensory disabilities emerged with the intent of providing a context of support and expertise.

By the early 1900s, compulsory education for students with disabilities was well under way in the United States. At this time, there was a strong belief that it was best to segregate so-called backward and feeble-minded children. One significant challenge concerning function that educational systems faced was identifying children in the upper echelon of those with disabilities who might benefit from academic instruction. Intelligence testing quickly took root as an instrumental resource to assess students' functioning. At that time, the majority of special education services were centered on students with types of severe physical, visible disabilities, and physicians led these services.

In the late 1800s and early 1900s, the negative attitudes toward people with disabilities held by those in the growing eugenics movement, especially attitudes toward those who are limited cognitively, competed with early advances in educating children with disabilities. Although in the 1930s, the eugenics movement did interrupt these advances, new social attitudes toward children with disabilities soon began to take root. These attitudes posited that every citizen had a right to instruction to help him or her develop into an asset to society. In addition, parents began to organize and insist that government and schools acknowledge and address the needs of children with disabilities. In general, services were aimed at overcoming children's impairments and involved segregated schools or classrooms.

In the mid-20th century, a philosophical shift sparked social changes that improved educational opportunities for youth with disabilities. By the 1950s, in response to dissatisfaction with the isolation of children with disabilities, the National Association for Retarded Children, a powerful parent advocacy organization headquartered in Minneapolis, began to work toward policy changes. During this time, individuals with emotional disabilities—and later, in the 1960s, people with intellectual disabilities—began to move out of institutions and into the community, thereby increasing demand for community-based supports and opportunities.

Not only were parent advocate groups and organizations central to the struggle for equal educational opportunities for children with disabilities, but federal legislation has also served a critical role in securing the civil right of children with disabilities to a free and appropriate education. In large part, these acts have focused on understanding the types and functions of disability. More recently, an emerging understanding of the social construction of disability has moved forward positive practices and greater appreciation concerning the way social structures can promote well-being and success among youth with disabilities. Providing free and appropriate education continues to be a challenge, but substantial progress has been made since federal legislation provided a public policy foundation for special education in the 1960s.

Although the rights of children with disabilities were attracting greater attention by the mid-20th century, many children with disabilities continued to be excluded from the public school system. Parents and guardians pursued litigation to secure the educational rights of children with disabilities. These lawsuits established the importance of better identification of children with disabilities and offering children with disabilities a state-supported education. Soon after these lawsuits, the Education Amendments of 1974 (Public Law [P.L.] 93–380) were passed, which included an amendment renamed the Education of the Handicapped Act (EHA) Amendments of 1975 (P.L. 94–142). These amendments guaranteed that a "free appropriate education," inclusive of special education and related services, be available to children with disabilities in "the least restrictive environment" (LRE; also known as the *least restrictive placement*). It created a time line for ensuring full educational opportunity for all children with disabilities and procedural safeguards for their identification, evaluation, and placement. It mandated that, when possible, educators place children with disabilities in regular classrooms and that testing and evaluation materials not be discriminatory. The act denied schools the ability to exclude children with disabilities. The EHA was later amended to include incentives for early intervention, special education programs in preschools, and transition services.

In 1990, amendments to the EHA were renamed and led to the creation of the Individuals with Disabilities Education Act (IDEA) (P.L. 101–476). These amendments mandated transition programs that had to be considered in a student's Individualized Education Program (IEP), which is a yearly plan of a child's educational goals and supports necessary to achieve those goals. IDEA also expanded many discretionary programs including services and research/dissemination centers for children with specific forms of disability. Under the amendments of 1997, IDEA emphasized not only physical access to education but also, to a greater degree, involvement in the general curriculum involvement and outcomes for children with disabilities.

IDEA was most recently reauthorized in 2004 (P.L. 108–446). The new regulations align IDEA with the No Child Left Behind Act of 2002 (P.L. 107–110) and include an increased focus on accountability to ensure that children with disabilities are held to high standards. These amendments also require policies and procedures intended to prevent inappropriate overidentification or misdiagnosis of minorities as disabled, identify children with disabilities in foster care, and involve parents in decision making.

Disability Prevention and Diagnosis

In order to provide and evaluate appropriate educational opportunities for youth with disabilities, this entry examines both how disability is prevented and how youth with disabilities are identified in education settings. In educational contexts, prevention of the development of a disability involves the identification of and treatment of students who do not have disability but display risk factors (e.g., limited oral language ability, difficulty counting) that increase the likelihood of a disability later in life. For example, children with developmental speech/language impairments are at higher risk for reading disability than peers with no history of speech/language impairment. Therefore, the early identification and treatment of reading deficits may aid in preventing the development of a disability among such students.

The diagnosis of a disability remains rooted in a medical model emphasizing type and, to a lesser extent, functioning. The identification of impairment in an individual is often coded using either the *International Classification of Diseases* (Ninth Revision, Clinical Modification) (ICD-9-CM) or the *Diagnostic and Statistical Manual of Mental Disorders* (Fourth Edition, Text Revision) (*DSM-IV-TR*). Individuals may be identified as having more than one impairment, including the primary diagnosis and any secondary conditions present. There are at least four important issues to consider in the diagnostic process: diagnostic methods, early identification, and two

diagnostic dilemmas—misdiagnosis and communicating effectively with parents.

Many disabilities are identified early in an infant's life, whereas others become evident when the child enters educational settings. Disabilities that may be diagnosed once the child is in educational settings include learning disabilities (LDs). *Learning disability* is a general term that describes a variety of specific kinds of learning problems. A learning disability may cause children to have trouble learning and using certain skills, including reading, writing, listening, speaking, reasoning, and mathematics. Learning disabilities are the most common disability found in children and adults: Almost 3 million school-aged children have some form of a learning disability and receive related special education services in school. More than half of all children who receive special education services have a learning disability. Approaches to the diagnosis of a disability vary depending on the nature of the disability; given the prevalence of learning disabilities among school-aged children, this entry describes the diagnostic process as an exemplar.

The most common approach to diagnosing an LD is a thorough, systematic evaluation of a child's strengths and weaknesses as a learner. These evaluative assessments may be written or oral, include an interview about the history and details of problems experienced, and are designed to identify IQ achievement discrepancies of students. However, the use of IQ achievement discrepancy procedures for the identification of children with learning disabilities has been persistently criticized because the diagnosis is based solely on the child's performance on IQ-related tasks and does not sufficiently consider other relevant factors. For example, such procedures do not take into account that environmental deficits, such as poor nutrition, may produce effects in a child's performance that mimic, rather than identify, a disability in a child.

Given the potential inaccuracies in identifying students with learning disabilities under this IQ-related method, alternative methods for identification and diagnosis have been offered, including response-to-instruction (RTI) models. RTI models are based on the premise that a student is identified as learning disabled when his or her response to an effective educational intervention is dramatically inferior to that of peers. Such models recognize that learning difficulty may lie within the child, within the instruction, or both. The models promote an earlier identification of LD, rather than waiting for students to show signs of failure, because the primary focus is on providing effective instruction and improving student outcomes. Diagnostic models that incorporate RTI characterize a shift in special education toward the goals of better achievement and behavioral outcomes for students identified with LDs, as well as those students at risk for LDs by heavily emphasizing both prevention and intervention.

During these diagnostic steps, early identification of students with disabilities to help them better understand and cope with their existing disability diagnosis is critical. Early detection is possible, and intervening earlier rather than later is most therapeutic and cost effective when addressing the needs and concerns of people with disabilities. Although there is a range of early intervention services that may be provided, they often involve physical, occupational, and/or speech therapy. Early intervention among those with LDs is important because the effects of learning disabilities tend to accumulate over time, and early detection can dramatically improve a child's life course and outcomes. As children fall behind in school, they may become more frustrated and increasingly feel like failures. Early detection is possible, and intervening earlier rather than later is most likely to be beneficial and cost effective when addressing the needs and concerns of people with disabilities.

Diagnostic Dilemmas

Currently, teachers in the United States determine whether a discrepancy exists between their expectations for a given student and the student's academic or social behavior in order to refer students to school psychologists for assessment of a disability. However, educators have been criticized for strategies that identify students of color, particularly those from low-income families, as disabled and place them in special education in the public school system. Nationwide, students of color are overrepresented in special education programs, indicating widespread misdiagnosis. Research suggests that some students of color are misdiagnosed with learning disabilities on the basis of biased assessments and assigned to remedial education programs that do not accurately address their learning needs. Overrepresentation in special education is most pronounced for Black children, who are nearly three times more likely to be labeled mentally retarded and nearly twice as likely to be labeled emotionally disturbed as White students in the United States.

There have also been criticisms regarding the identification and treatment of disabilities among students who are English-language learners. Students designated as limited in English proficiency are those who have received classroom instruction in English for less than 3 years and whom school personnel deem incapable of participating in statewide assessments in English. However, students born outside of the United States who have limited English proficiency are often inappropriately placed in special education programs based on tests given in English while in the United States. As a result, students with limited English proficiency are also disproportionately represented among students with disabilities.

Given the propensity of misdiagnoses, some experts believe identifying and diagnosing students with learning disabilities should be more holistic. Conscientious diagnostic personnel need to have an in-depth understanding of each learner and an awareness of the multiple competing factors in addition to a disability-related impairment that may give rise to lower test performance. They understand the right of every child to have an appropriate education in the least restrictive environment. The absence of these critical elements may jeopardize the effectiveness of both the prevention and identification of disabilities and of interventions for students with disabilities and their families.

Communicating a diagnosis of disability to parents can be challenging not only because of possible misdiagnosis, but also because of the process of accepting and understanding the disability and the ensuing need for modifications in rearing, educating, and advocating for a child that may be necessary. Parents' responses to having a child diagnosed with a disability may be diverse, both positive and negative. Timing; privacy in communicating; obtaining clear information about disability type, function, and social construction; continued and consistent contact with disability professionals; supportive counseling, and the understanding and support of family, friends, and school staff are important to help parents cope with the discovery of their child's disability. Once a disability has been identified, it is important to involve parents in managing the disability with their children. They are usually able to provide more ongoing support than others, and they often appreciate learning practical skills to help their children. Many parents report that, despite the considerable and sustained stress involved in parenting a child with a disability, their experiences have been personally and positively transformative.

Educating Youth With Disabilities

Upon a student's diagnosis with a disability, the school system needs to provide an appropriate education given the specific type of disability and learning needs of the student. Early intervention typically serves children from birth to 3 years old who have been identified as having a disability. Often, parents work closely with schools or other agencies to form an Individualized Family Service Plan (IFSP). From age 3 up until ages 18 to 21, the school system and/or other agencies provide educational and related support services as specified in the IEP. Given the diversity across disabilities, the ways in which youth with disabilities are educated and supported must vary accordingly. Interventions are generally specific to the disability and abilities of the student, such as extensive training for students who have limited reading skills. Intervention for a disability may be provided in a number of ways by teachers, peers, and/or parents.

Federal IDEA legislation requires that students with disabilities be placed in the least restrictive environment (LRE); that is, to the extent possible, students with disabilities are educated with students without disabilities. As such, many youth with disabilities are educated primarily by general education teachers. In general, teachers are expected to identify effective ways to meet the needs of a diverse group of students and ensure their success in the classroom. Regular education teachers often vary in their willingness, resources, leadership support, and preparation to make adaptations to their curriculum or teaching techniques to accommodate students with disabilities.

As educators place youth with disabilities in LREs, attention to both their physical and social integration is critical. Physical integration is the placement of students with disabilities in traditional learning classrooms. Social integration relates to facilitating opportunities for students with disabilities to build relationships with nondisabled peers and general education teachers. Physical integration into traditional classrooms is a precursor to social integration. Students with disabilities need to be learning in traditional classrooms where social interactions with nondisabled students and general education teachers can occur routinely before they can have the social opportunity to build connections with these peers and

teachers. Cooperative learning interactions between students with and without disabilities may promote an increase in supportive behaviors in students without disabilities to their disabled peers as well as greater frequencies of friendship and social activities outside of the classroom.

Determining appropriate classroom accommodations that a student may need given a specific disability is also a critical aspect to educating youth with disabilities. Outside of necessary accommodations to the physical environment, such as space for a wheelchair and room to maneuver it so that a student can participate fully in class activities, school and educational psychologists often direct their attention to the area of classroom instruction and testing accommodations.

The recognition that students with disabilities may need individualized instruction has also been applied to assessment, leading to individualized types of assessment and analysis of underlying processes thought to relate to the specific disability. Typically, learning assessment refers to state-mandated testing to determine student performance. Uniform performance evaluations present challenges to students with disabilities. Recently, constructive trends in assessment have shifted toward a greater use of assessments that are specific to a given subject matter, have curricular validity, account for contextual influences, and are developed from models of cognition. However, because school systems are often outcome based, minimum competency requirements apply to students with disabilities. Often, a marked discrepancy may exist between the curriculum taught and the content of standardized tests. Consequently, learning assessment and education of students with disabilities have many challenging issues that remain unresolved. Schools seek to understand and educate individuals with disabilities effectively and often distinctively in a context of limited resources, systemwide norms, and legal mandates.

School systems are also focusing on how to best prepare youth with disabilities for successful adult lives, including higher education, independent living, and competitive employment. This increased focus on transition has included attention to how to best involve parents in the transition process. Although IDEA has mandated since 1997 that transition planning begin at age 14, implementation has not been complete. Transitions plans, included in students' IEPs, usually focus on function and address the students' long-term goals, including employment preparation, technical training, or educational course plan. As such, educators need to be knowledgeable about each individual student's abilities and identify appropriate opportunities to best facilitate the goals outlined in the transition plan.

Educational Psychology and Students With Disabilities

Educational psychologists may determine how to administer assessments of students with disabilities and evaluate their results. They are able to educate teachers about appropriate types of instructional techniques and curricula for students with differing disabilities and abilities. Furthermore, drawing on existing studies and conducting research, they can adapt information and techniques used by special education teachers to be increasingly practical for a diverse classroom of learners. Educational psychologists are also often involved in evaluating the effectiveness of a specific intervention to determine whether the outcomes are appropriate and beneficial.

Educational psychologists can also facilitate successful transitions for students with disabilities by addressing their needs through increasing access, skill building and enhancement, and preparation for the transition ahead prior to their graduation from high school. They have been instrumental in developing transition programs that are effective and efficient for students with disabilities after high school, such as competence enhancement, job training, school-to-work programs, or college-preparatory programs.

In conclusion, the education of students with disabilities is one of the major challenges facing contemporary society. The historical record is marked by positive initiatives, good intentions gone awry, and malicious actions and policies. Over the past generation, society has come to recognize the potential of people with disabilities and has safeguarded these individuals to a much more significant degree than previously in law, policy, and practice. Schools are making a stronger effort than ever before—albeit partial, conflicted, and imperfect—to assess, educate, and transition children and youth with disabilities to positive futures. Using type, function, and social construction perspectives of disability, educational psychologists can enable schools to meet this challenge and enable students with disabilities to realize their full range of abilities and make valuable contributions to society.

Christopher B. Keys, Katherine E. McDonald, Shannon Myrick, and Terrinieka T. Williams

See also Early Intervention Programs; Inclusion; Individuals with Disabilities Education Act; Learning Disabilities; Least Restrictive Placement; Mainstreaming; Special Education

Further Readings

Accardo, P. J. (1996). *The invisible disability: Understanding learning disabilities in the context of health and education.* Washington, DC: National Health & Education Consortium.

Harry, B., & Klingner, J. K. (2006). *Why are so many minority students in special education? Understanding race and disability in schools.* New York: Teachers College Press.

Keogh, B. K., & MacMillan, D. L. (1996). Exceptionality. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 311–330). New York: Macmillan.

National Information Center for Children and Youth with Disabilities. (n.d.). *IDEA related publications.* Retrieved May 17, 2007, from http://www.nichcy.org/ideapubs.asp

Piercy, M., Wilton, K., & Townsend, M. (2002). Promoting the social acceptance of young children with moderate-severe intellectual disabilities using cooperative-learning techniques. *American Journal on Mental Retardation, 107*, 352–360.

U.S. Department of Education. (n.d.). *Building the legacy: IDEA 2004.* Retrieved May 17, 2007, from http://idea.ed.gov/

Vaughn, S., Bos, C. S., & Schumm, J. S. (2000). *Teaching exceptional, diverse, and at-risk students in the general education classroom* (2nd ed.). Boston: Allyn & Bacon.

World Health Organization. (2001). *International classification of functioning, disabilities and health.* Geneva: Author.

DISCIPLINE

One of the results of the heightened concern about school safety is increased attention to school discipline as a factor in ensuring safe and orderly schools. Disruptive behavior in schools not only poses safety problems when incidents involve weapon possession, violence, or substance use, but also interferes with instructional efforts by teachers and with learning conditions for students. The most common understanding of the term *school discipline* involves punishment of student misbehavior by removal from the classroom or the school (i.e., office referrals, suspensions, and expulsions). School removal rates have been used as an indicator of safe and orderly school campuses. In the 1999 National Center for Education Statistics report, a high association was found between principal perceptions of discipline problems and school crime statistics. Within and between districts, higher numbers of disciplinary incidents are associated with higher rates of misbehavior. As well, there is an assumption that fighting and aggressive behavior can escalate into more violent events, and even nonviolent forms of student misbehavior can lead to unsafe school environments. High numbers of suspensions also have been associated with negative academic indicators such as grade retention, dropping out of school, alienation from school, juvenile delinquency, and drug use.

More recently, attention has been drawn to so-called low-level violence, or incidents of behavior, such as bullying, peer sexual harassment, and victimization. Chronic bullying has been characterized as a contributor to disturbed mental status and to the potential for retaliatory, aggressive, and perhaps violent behavior. As a result of greater awareness of the negative impact of incidents of low-level violence, such behaviors have been added to lists of suspendable offenses, along with more physical forms of threat and aggression. Zero-tolerance policies have taken disciplinary policies to an extreme, broadening the scope of exclusion to behavior that, although related, may not be associated with greater likelihood of violence and disorder (e.g., plastic knifes, ibuprofen tablets—interpreted as a "substance" being abused). These changes in school discipline policy are indicative of concern about behaviors that threaten psychological and developmental safety as well as physical safety.

Despite the recent association of school discipline with safe and orderly schools, historically, school discipline has included a broader range of practices that includes prevention of misbehavior, remediation of behavioral problems, and exclusion as a punishment for severe forms of misbehavior. The Latin root of the word *discipline* is actually from the verb "to learn." Thus, the broader conception of school discipline as an opportunity to teach positive behavior is in keeping with the roots of the terminology. This entry reviews the most recent trends in school discipline: It describes the process of zero tolerance as reflected in the indices of office referrals, suspensions, and expulsions; describes which students are being disciplined; and reviews best practices for implementation of

discipline practices that focus on the broader conception of discipline as an opportunity to learn.

School Removal: Rates and Groups

The Gun Free Schools Act of 1994 required states to have legislation to mandate the expulsion of youths who bring guns to school. As safe and orderly schools became of interest to this nation's educators, attention turned to gathering information about the rates of utilization of disciplinary measures. Research noted a dramatic increase in the use of exclusions in early and late 1990s. Nationally, whereas less than 4% of students across the nation were suspended or expelled in the 1970s, the 1990s saw those rates increase to 8% to 10%.

As these rates climbed, interest turned to the characteristics of this increasing population of excluded students. Students who find themselves involved in disciplinary incidents at school come to that place from a variety of circumstances. Students who are unlikely to pose a serious threat of danger to the school environment commit the majority of the offenses. They include a diverse group of students who take different pathways to the expulsion event and have strikingly different etiologies for their behavior. Not all students have a chronic history of discipline problems. Some come to the expulsion event through attendance problems, school disconnection, involvement in drugs, or a chronic history of emotional difficulties. Notably, however, students recommended for expulsion generally perform well below average in terms of grades and achievement scores. Thus, the one common red flag is poor school performance.

Two groups of students are particularly vulnerable to discipline under zero-tolerance policies: underrepresented minorities and students with disabilities. Data about suspensions and expulsions are consistent in the characterization of the overrepresentation of minority students. In particular, rates for African American students have been commonly reported as triple their representation in the student population.

Investigations reveal that African American and Latino students were more likely to be referred for discipline and more likely to be disciplined. Disciplinary actions were taken for these groups for minor misconduct, with the punishments seemingly out of proportion to the offenses. For example, discretionary offenses such as "defiance of authority" and "disrespect of authority" are more likely to be treated as exclusionary offenses with these populations.

Rates for students with disabilities have indicated double and sometimes triple their expected representation in the groups of students involved in disciplinary incidents. Special education students who find themselves in trouble at school are likely to have documented emotional, familial, or behavioral risk indicators that parallel their special education histories. Particularly salient risks for students include a history of attention deficit hyperactivity disorder (ADHD) or conduct problems, chronic or crisis-related family problems, and a history of inconsistent special education placements and interventions.

School Discipline as an Interaction Between Student and School Characteristics

Although general trends in the incidence and use of school disciplinary processes and general characteristics of students who are involved with these processes can be identified, there is great variation between schools in terms of how many students are disciplined and who gets disciplined for what behaviors. Student developmental and behavioral trajectories are inextricably tied to schooling practices. Appropriate practices may serve to reduce behavioral problems on a school campus, whereas lack of consistent and appropriate programming, education, and treatment may exacerbate developmental problems experienced by a student.

In addition to student and school environment interactions, student involvement in disciplinary events is not simply a one-time event; often, the interactions described above occur over time. There are risks and resilience (protective) factors within students and within their school environments that *precede* problem behavior leading to disciplinary actions. In the context of (*during*) delivering the consequences (office referral, suspension, expulsion), there are practices that improve or exacerbate the situation. Finally, intervention *after* the incident may cause additional risk or afford protection and positive development. Using the characterization of school discipline as a process (before, during, after) rather than an event and the framework of risk and resilience at the different stages of the process, school discipline and school removal will be described further in the following

stages: prevention, application of the consequence, and intervention.

Prevention:
Before an Incident Occurs

In recognition that student involvement in school discipline is affected by the context and environment of the schooling experience, an exploration of the philosophical orientation toward student discipline and student behavior is instructive. Despite the origins of the term *discipline* in a learning framework, student misbehavior has been viewed as a reason for excluding a student from school.

There is debate about whether or not schools are obligated to educate students who egregiously and continually violate school rules. For example, social maladjustment does not qualify for the protections and guarantees of free and appropriate education under special education law, despite the fact that these youths may present the greatest long-term social and economic costs to society. The assumption remains that somehow these students should lose their right to a public education if they cannot abide by the behavioral parameters set by schools. This social policy conflict may be result of the competing worldviews representing a "constrained" view (humans are imperfect; they misbehave, are responsible for that misbehavior, and should pay the consequences) and an "unconstrained" view (humans are perfectible; they misbehave due to special causes, ignorance, or social inequities; special circumstances require individualized application of consequences). In the constrained view, the obligation to educate a student ends after the behavioral transgression; in the unconstrained view, it does not. In the constrained view, expulsion is an event that ends the educational right; in the unconstrained view, the educational process continues, albeit with changes to accommodate the needs and behaviors of the student. The constrained worldview aligns with zero-tolerance policies, where little flexibility exists in applying the consequence or in considering the history and characteristics of the student who commits the offense. The unconstrained worldview involves a broader view of the influences on child behavior. This worldview supports the creation of educational and personal-social interventions to develop positive behavior rather than depend on punitive interventions. Students in schools where discipline practices are guided exclusively by the

constrained worldview are at greater risk of being excluded from school, whereas the unconstrained worldview is more likely to lead to "protective" discipline practices.

Preventing Misbehavior
at the School Level

The presence or absence of effective prevention programs that serve students with a range of risks and abilities is likely to affect the rates of school disciplinary events. Proactive school discipline programs have demonstrated effectiveness in reducing student behavior problems. The existence of a schoolwide discipline plan is a key strategy used in schools for preventing behavior problems that lead to student exclusion from school. Such a plan includes a clear statement of rules and expectations, consistently communicated and applied consequences for rule-breaking behavior, concrete efforts to teach students appropriate behavior, and positive consequences available for positive behavior. With this system in place and consistent implementation by all staff, students have a better chance of behaving in ways that will maximize their inclusion into the activities of the school.

The following essential characteristics of schools with low school disciplinary referrals have been identified: (a) Administrators and teachers demonstrate ownership of discipline-related problems that students present; (b) opportunities exist to develop strong bonds between teachers and students; (c) ongoing staff development is conducted about best practices in handling student misconduct; (d) a schoolwide code of conduct and expectations is promoted; (e) a process, location, and plan for students who need to "cool off" is available so that more severe outbursts can be prevented; (f) student sanctions are considered on a case-by-case basis with input from students and parents; (g) community participants are welcomed into the school, including parents, mental health and juvenile justice professionals, business leaders, and so on; (h) explicit efforts are made to show students that they are valued and respected members of the school community and that they are expected to uphold high behavioral and academic standards; (i) a wide variety of prevention and intervention support programs are available for students; and (j) the school physical environment is a friendly and welcoming space. These school characteristics add to the likelihood that students will stay within the behavioral standards of

the school and not be exposed to school removal actions.

Preventing Misbehavior at the Classroom Level

The classroom is one of the main contexts where a student's behavior comes into conflict with the rules and norms of the school. The teacher's ability to handle misbehavior and encourage positive behavior affects student behavior. A student who struggles with behavior issues in the classroom is at greater risk if a teacher has few strategies for addressing developmental lags and uses increased negative attention in an effort to control behavior. Teachers vary in their tolerance of and their ability to handle student misbehavior in their classroom. It is not uncommon to have a high number of referrals coming from a small proportion of teachers on a school campus. Classrooms that are characterized by low rates of academic engagement, praise, and reinforcement and that have high rates of reprimand are associated with high rates of misbehavior and show a cycle of negative student-teacher interactions.

In contrast, teacher management strategies, effective instructional techniques (e.g., class-wide peer tutoring), early intervention for students with learning problems, and positive teacher–student relationships are critical components for keeping students in the classroom or avoiding the necessity of excluding them for disciplinary reasons. Effective instruction and positive student–teacher relationships should be accompanied by opportunities for students to be involved in activities to promote the development of desired social and emotional skills.

The protective or preventive actions that are available to educators to help reduce misbehavior at school require strategies that are specifically aimed at teaching social skills and implementing behavioral strategies. Attention to building social skills and positive behavior may seem at odds with the current climate of academic accountability that is evident in this nation's schools. Time spent on direct student interventions or staff training on social and behavioral intervention may not get as much support or attention as similar programming for academics. Furthermore, some reports indicate that academic sanctions have made their way into school disciplinary logs (e.g., student detention for not finishing schoolwork or coming prepared to class). This situation reiterates the importance of considering schooling policies and trends as contributors to levels of school discipline utilization.

Protective Factors in the Discipline Process

Office Referrals

If a school has a graduated discipline process in place, the student trajectory through this process might look something like this. The student experiences several attempts by the teacher to keep him or her in the classroom. If these attempts do not work, an office referral may result. Risks to students at the office referral stage exist when there is no systematic agreement about behavioral expectations within classrooms and throughout the school. There may be wide variations from teacher to teacher about what constitutes behavioral transgressions. Whereas some teachers may refer many students for small transgressions, some may refer few and use the referral for significant rule-breaking behavior. One often-ignored risk occurs when any one student is referred many times and spends a significant amount of time away from the classroom, thereby missing important academic engagement time. Another risk is when the intervention at the office or principal level is ineffective. The student thus not only misses academic time, but also learns nothing from the experience.

Protective factors may be built into the office referral system. Schoolwide agreement on what constitutes an office referral, systematic collection of data to determine the incidence of student behaviors, and examination of teacher referral patterns are practices that allow for periodic examination of the effectiveness of existing behavioral interventions. Examination of who gets referred for what is likely to reveal that a small group of students is involved in the majority of referrals. The intervention for these students should then be intensified. Early, effective intervention for misbehavior serves as a preventive mechanism for continued trajectories toward serious behavioral problems.

Suspension

Suspension from school is a more serious consequence that is regulated by state and local statute. Most states, as guided by the Gun Free Schools Act, mandate suspension and expulsion for possession of a weapon (firearms or bombs) at school. States recently

have added offenses to this list that include threats, sexual harassment, and bullying. Although suspension is required for the most serious offenses, discretion of the school administrator is available for lesser offenses such as school disruption, defiance of authority, attendance problems, tardiness, inappropriate dress, and so on. It is in the areas of discretion where possibilities for risk and protection exist. Zero-tolerance approaches to suspension, even for minor offenses, lead to higher levels of suspension, which are associated with lowered school effectiveness. One reason for this association is that suspension as an educational strategy has not been proven effective. Whereas school exclusion of students who pose a danger to the school makes sense for achievement of safe and orderly schools, suspension for lower-level offenses, as with office referrals, leads to reduced academic engagement time and may not lead to the desired change in future behavior. Automatic exclusions, in the absence of exploration of the root causes of misbehavior or without intervention designed to specifically address the problem behavior, contribute to limited effectiveness of the suspension.

In order to avoid overuse of suspension as an ineffective consequence for misbehavior, it should be used thoughtfully and judiciously. Data on who gets suspended for what offenses should be carefully monitored. As with office referrals, repeat offenders account for a good proportion of the suspensions, and appropriate interventions can be targeted to this group. Schools now use in-school suspension in order to monitor out-of-class time and lessen the impact of loss of academic engagement time.

Expulsion

Suspension is the first step toward expulsion. School administrators must recommend expulsion for some offenses. The balance that school administrators must face in final expulsion from school is maintaining the safety of their school campus and denying the educational opportunity of the student being excluded. The risks for the student are those associated with violations of due process and bias that may enter the process, resulting in an unfair exclusion from school. The factors associated with how students arrive at the expulsion event may be framed as risk factors; that is, students who get expelled demonstrate risk in areas of ethnicity, disability, school maladjustment, lack of parent advocacy, and alienation from activities and people at school.

However, there are protections embedded within the school expulsion process. The final decision about school exclusion is a school board decision. The board may have the discretion to suspend the expulsion order and return the student to the district, perhaps to another school within the district. Furthermore, many states now require that alternative education be provided to students upon expulsion. Another positive influence on a district action to stop short of expulsion is the presence of an advocate in the student's school, community, or family. These individuals might vouch for the overall character of the student, ensure that due process is being followed, or advocate for the continuity in current educational and therapeutic interventions. A student who has multiple positive factors in his or her school record (e.g., extramural activities, leadership roles, good grades) is less likely to receive the full expulsion consequence. Finally, special education services and support serve as a deterrent to school expulsion (i.e., beyond the initial recommendation), especially given the protections provided to special education students under federal law.

Continued examination of school expulsion data, as with office referral and suspension indices, is recommended in order to provide a check on who is being expelled and what outcomes befall students who are expelled. There is a dearth of information about outcomes for expelled students in the research literature to date.

Use of School Discipline Indices

A cautionary note about disciplinary indices is offered here. Despite attempts to associate group- and individual-level statistics with negative, risky outcomes, these indices are not necessarily intended to serve as reliable and valid measures of misbehavior. Rather, they have been developed as legal and record-keeping measures. Office referrals help communicate between teacher and principal about student misbehavior. Suspension and expulsion data result from legal due process procedures for excluding students from school. Because these indices were not necessarily intended to be used for group and individual description and prediction, these systems are fraught with error in design and implementation. This lack of reliability of office referral, suspension, and expulsion counts compounds the basic challenge of finding evidence that these processes are effective. The increased use of suspension/expulsion is especially

contradictory within an educational/political climate that demands utilization of evidence-based practice.

Intervention After Disciplinary Actions

A student may receive suspension from school and never get in trouble at school again. However, students often continue to reoffend and need more intensive intervention for their behavior and its underlying roots. As an alternative to school removal for recurring misbehavior, the most common interventions include targeted interventions for misbehavior and aggression and placement in alternative settings.

Targeted Interventions

Targeted interventions may include intervention for the student's problem behavior (anger management, conflict resolution, social skill instruction) as well as intervention in the nonschool settings in the student's life (home, community). FAST Track is an example of a comprehensive program that includes intervention in the student's home (parent training) as well as in school (academic tutoring). The Multisystemic Therapy (MST) program is an evidence-based program that includes intensive, therapeutic intervention within the student's home environment. Another targeted intervention receiving attention is restorative justice. A restorative justice program involves the offender and victim in a process of identification of the offense and attempting to repair the damage. This type of program facilitates the building of relationships and sense of community as opposed to punishment-based exclusion.

Alternative Settings

Alternative education programs may provide services to students who have been expelled, but not all students who are expelled are afforded the right to attend an alternative education program. There are no federal laws mandating that all states provide alternative education placement for students who are expelled. Currently, the decision as to whether these alternative programs are discretionary or mandatory in nature varies state by state.

Perhaps because of the wide variety of students served, as well as the differing legislation across states regarding implementation of programs, very little is known about how well alternative education placements serve students, particularly those who have been expelled. Some programs have made claims that their alternative education programs have had some success in (a) providing GED completion, remedial assistance, and vocational training; (b) developing communication, coping, and self-control skills; and (c) keeping students in school. Strategies that are likely to lead to successful alternative education programs include the following: (a) low student–teacher ratio, (b) highly structured classrooms with behavioral classroom management, (c) positive rather than punitive emphasis in behavior management, (d) availability of adult mentors, (e) individualized behavioral interventions based on functional behavioral assessment, (f) social skill instruction, (g) high-quality academic instruction, and (h) parent involvement.

Involvement in either targeted interventions or in alternative education provides students who have been excluded from school the protective opportunity to gain skills that will be necessary to re-engage in the school environment in a productive way. Students who do not receive programming to build skills and change behavior are at high risk for failing again upon their return to the school environment.

Best Practices

To summarize the basic principles of best practices in school discipline, the American Bar Association passed a resolution in which it outlined three principles that apply to the protective processes described in this entry. The Association opposed schools' zero-tolerance disciplinary policies that fail to take into account the circumstances or nature of an offense or an accused student's history. It supported three principles concerning schools discipline: (1) Schools should have strong policies against gun possession and be safe places for students to learn and develop; (2) in cases involving alleged student misbehavior, school officials should exercise sound discretion that is consistent with principles of due process and consider the individual student and the particular circumstances of misconduct; and (3) alternatives to expulsion or referral for prosecution should be developed that will improve student behavior and school climate without making schools dangerous. Effective practices associated with these principles will keep students in schools; ineffective practices will create a great risk for student removal from school.

Gale M. Morrison

Further Readings

Bear, G. G., Cavalier, A. R., & Manning, M. M. (2002). Best practices in school discipline. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology IV* (pp. 977–991). Bethesda, MD: National Association of School Psychologists.

Furlong, M. J., Morrison, G. M., Cornell, D., & Shiba, R. (2004). Methodological and measurement issues in school violence research: Moving beyond the social problem era [Special issue]. *Journal of School Violence*, *3*.

Skiba, R. J., & Noam, G. G. (Eds.). (2002). *Zero tolerance: Can suspension and expulsion keep school safe?* San Francisco: Jossey-Bass.

Skiba, R., & Rausch, M. K. (2006). School disciplinary systems: Alternatives to suspension and expulsion. In G. G. Bear & K. M. Minke (Eds.), *Children's needs III: Development, prevention, and intervention* (pp. 87–102). Washington, DC: National Association of School Psychologists.

Skiba, R., Reynolds, C. R., Graham, S., Sheras, P., Conoley, J. C., & Garcia-Vazquez, E. (2006). *Are zero tolerance policies effective in the schools? An evidentiary review and recommendations.* Washington, DC: American Psychological Association.

Sugai, G., Horner, R. H., Sailor, W., Dunlap, G., Eber, L., Lewis, T., Kincaid, D., Scott, T., Barrett, S., Algozzine, R., Putnam, R., Massanari, C., & Nelson, M. (2005). *School-wide positive behavior support: Implementers' blueprint and self-assessment.* Eugene: University of Oregon Press.

DISCOVERY LEARNING

Issues concerning human learning are among the critical topics in educational psychology, child development, and cognitive science. One central focus has been on the issue of how students learn and teachers teach best, and discovery learning versus direct instruction has been a contentious debate in modern educational theory and practice. Stemming from the theoretical perspective of constructivism, discovery learning is believed to be a tool for facilitating the creation and organization of knowledge, as well as the transfer of that knowledge across different contexts. This approach contrasts with views that emphasize direct instruction from teacher to student. This entry addresses the following central issues: how discovery learning is defined, the empirical evidence in favor of discovery learning or direct instruction, and the facilitation of discovery and transfer.

Definition

Discovery learning is a general approach that involves mindful participation and active inquiry in the acquisition of concepts and strategies. In classroom contexts, it refers to a form of curriculum in which students are encouraged to actively explore and figure out the concepts, solutions, or strategies at hand. A widely accepted idea is that discovery learning is the most appropriate and effective approach to facilitating deep and lasting understanding. This approach is often contrasted with direct instruction or expository learning, which typically refers to traditional, content-oriented methods whereby the instructor lectures to students. Learning associated with direct instruction is often believed to be less engaged and less active, and thus less effective.

The spirit of discovery learning can be traced back to philosophical traditions. The French philosopher Jean-Jacques Rousseau believed that children do not learn well via instruction and therefore should be given maximum freedom to explore their surroundings. He claimed that children learn primarily from spontaneous exploration of the environment and from interactions with people and objects. John Dewey posited that the essence of education is learning by doing and exploring. According to Dewey, discovery learning provides the "intimate and necessary relation between the processes of actual experience and education."

Discovery learning is at the core of the constructionist view of education. According to Jean Piaget, children learn by interacting with the environment, both physically and mentally. He emphasized the importance of "cultivating the experimental mind." Educators must teach children how to use discovery as a tool for constructing and acting on their worlds. Another constructionist, Jerome Bruner, was an early advocate of discovery learning environments. He believed that allowing students to learn by constructing knowledge structures would lead to improved intellectual ingenuity and persistence, as well as increased motivation to learn. He claimed that

emphasis on discovery in learning has precisely the effect on the learner of leading him to be a constructionist, to organize what he is encountering in a manner

not only designed to discover regularity and related-ness, but also to avoid the kind of information drift that fails to keep account of the uses to which information might have to be put (Bruner, 1962, p. 87).

Thus, discovery learning is effective at helping the student uncover the underlying principles of a domain and in generalizing that knowledge to other tasks and contexts. Despite the widespread appeal of discovery learning and strong endorsements by a wide range of theorists and educational researchers, major issues and controversies remain regarding the role of discovery in learning and the evidence supporting this approach.

Empirical Evidence

Over the past half century, psychological researchers and educators have engaged in substantial debate about which approach, discovery versus expository learning, is more appropriate and effective. Although many have favored discovery learning over expository learning, others argued against discovery learning. More recent empirical research has generated mixed findings supporting each view. In particular, there are two related issues concerning the role of discovery or direct instruction in learning outcomes: (1) Can students *discover* strategies or rules? If so, (2) can they *transfer* what has been discovered to other contexts and situations?

The first issue is whether students are capable of making discoveries on their own, rather than being taught. There are ample examples of students' discovery of problem-solving rules or strategies. For example, even 5- and 6-year-old children are capable of discovering new rules for solving problems on the balance scale task. This task involves making predictions about whether the scale will go down on one side or stay level when a set of weights is placed on each side of the scale. How the scale moves depends on two factors: the number of weights placed on each side and the distance of the weights from the fulcrum of the scale. Children at this age typically make predictions based on the amount of weight on each side, but not distance. However, after having an opportunity to make predictions and then observe how the scale moves, children discovered the relevance of distance to the task. Therefore, children discovered a more advanced problem-solving rule in the sense that the experimenter never explicitly told children to pay attention to the dimension of distance.

Still, the effectiveness with which students discover rules or strategies depends on the mental activities in which the students engage and the guidance that teachers provide during the learning process. Discovery learning ranges from *pure discovery* to *guided discovery*. For the former type of learning, the learner explores the problem area in an unconstrained environment without direction or feedback from the teacher. In contrast, the latter type of learning involves supervision and guidance during the discovery process. In a 2004 article in *American Psychologist*, Richard Mayer reviewed studies of discovery learning between the 1960s and 1980s and argued that evidence in support of guided discovery was more prominent than for pure discovery. The findings of discovery learning of problem-solving rules, conservation strategies, and computer programming concepts all indicated that pure discovery is much less effective, as the learner often fails to recognize key principles and strategies and thus is not able to assimilate new experiences with existing knowledge. In contrast, guided discovery has proven to be more effective than pure discovery in promoting learning and transfer. Learners benefit from scaffolded experience, whereby the instructor provides constrained problem spaces to explore, encourages cognitive activities, or redirects learners when they went astray.

Another related issue is whether learners transfer more effectively when rules or strategies are discovered or when they are taught directly. Several studies have demonstrated that students who discover new concepts or strategies are more likely to extend this knowledge to new tasks than those students who learn from direct instruction. For example, in a 1990 paper in *Cognition and Instruction*, Mark McDaniel and Mark Schlager explored college students' strategy learning in solving Luchin's Water Jar problems. In this task, the solver must use a given array of three jars (A, B, C) to measure out a particular amount of water into a larger jar. Solution of the problem typically requires the use of an equation based on a combination of the three jars (e.g., $1A - 1C - 2B$). It was found that requiring learners to discover a solution strategy for this problem led to better performance on a transfer problem, which required the generation of a new solution equation. Thus, requiring students to generate equations based on the given tools proved to be most critical, and discovery conferred benefits beyond instruction.

In contrast, other experimental research comparing learning and transfer outcomes between instruction

and discovery suggests otherwise. In a 2004 *Psychological Science* article, David Klahr and Milena Nigam compared elementary school children's learning and transfer of the Control of Variable strategy in experimentation. They assessed children's conceptual understanding that only unconfounded experiments generate clear experimental evidence. Specifically, children were asked to design appropriate experiments based on the manipulation of appropriate variables and control of potentially confounding variables. In this study, one group of children was taught the strategy through direct instruction—an experimenter explained the importance of controlling confounding variables and manipulating variables of interest. A second group of children was allowed to interact freely with the problem materials. Therefore, if the children learned the strategy, it was through a process of discovery. The results showed that more children in the instruction condition learned the strategy compared to the children in the discovery group. Furthermore, many more children in the instruction condition were able to transfer what they learned. Numerous other lab experiments and classroom studies have also demonstrated the importance of instructional support and intervention during learning of mathematical and scientific concepts.

These mixed results suggest that it might not be productive to simply classify learning as discovery based or instruction based. Whether children discover target rules or strategies and how they transfer what they learn to other tasks depend on children's ages, their initial knowledge, the nature of the tasks, and the cognitive activities engaged during the process of learning. These seemingly contradictory findings, in fact, point to the need to recognize various forms and types of discovery learning between the two extremes of the spectrum: direct instruction and pure discovery. Perhaps the key is how to provide guidance during discovery. Psychological studies have pointed to several promising approaches to promoting discovery and transfer.

Facilitating Discovery and Transfer

Task Selection

A teacher's guidance can play a critical role in supporting discovery when learning materials are carefully organized and presented so as to streamline the discovery process. For example, in a 1976 study published in *Cognitive Psychology*, Siegler provided some

5-year-old children with feedback from problems in which the weights were the same on both sides but were placed on different pegs from the fulcrum. Other children received feedback from more complex problems in which one side had more weight, and the other side had the weight farther from the fulcrum. The first group learned more effectively and discovered the role of distance, whereas the second group did not learn to incorporate the distance dimension into their predictions. This moderate-discrepancy hypothesis suggests that the most effective experiences for promoting discovery are those that are somewhat, but not greatly, beyond children's existing problem-solving abilities. The role of guidance in organizing tasks for facilitating discovery is also evident in a study with college students by Zhe Chen and Mo Lei. Students who received more diverse instances of Luchin's Water Jar problems (diverse equations such as $1A - 2B + 1C$ and $1A - 2C$ and different types of problems such as volume, length, and area problems) were more likely to formulate a general problem-solving principle and apply it to novel tasks than these who received very similar tasks.

Guiding Questions and Self-Explanations

Another approach for effectively guiding discovery is the use of probing questions to direct students' attention to important features of the problem. In the case of learning the Control of Variable strategy, Chen and David Klahr conducted a study, published in 1999 in *Child Development*, in which they compared and contrasted two learning conditions in which elementary school children learned the Control of Variable strategy. The children's task was to design a series of valid experiments. In each, a single contrast was made (e.g., setting up two ramps that differ only in steepness, but not in other variables, in order to test the effects of steepness on how far a ball would travel down on the ramp). Although no explicit rationales were provided in either condition, two learning conditions differed in whether the experimenter asked children to explain why they designed a particular test and how they would interpret the experimental outcomes. Children in the directed questioning, or probe, condition thus had the opportunity to generate explanations of their reasoning behind their choice of objects for their experiment and for the possible conclusions from their design. The results indicate that elementary school children benefit from the systematic questions designed to promote deeper understanding. In contrast, allowing

a child to interact freely with experimental materials/ devices without asking guiding questions is insufficient for promoting discovery and transfer. Children's superior performance in learning and transferring the Control of Variable strategy in the probe condition further supports a growing body of studies demonstrating that the opportunity to generate self-explanations enhances children's learning.

Self-Generations

A number of studies have demonstrated the effects of mental activities during learning on discovery. Requesting students to generate evidence for their own opinions or to provide instances of principles has proven effective for fostering deeper understanding. Earlier research suggests that beginning lessons by asking students to generate their own ideas about phenomena, instead of directly telling students correct answers, is effective in facilitating formula comprehension and solving transfer problems. Similarly, encouraging students to generate concrete examples of abstract principles has been found to enhance understanding of the concepts and facilitate subsequent transfer. Other recent studies have shown that when students are asked to recall similar information, they tend to use superficial information in solving problems. In contrast, when students are required to process information by generating analogies, they tend to use underlying structural information.

Future Research

Discovery learning has long been favored as an effective approach to acquiring concepts, rules, and strategies. Despite the lasting and widespread appeal of active, mindful, explorative, and inquiring-oriented learning, the concept of discovery learning has been more of a philosophical belief or educational ideology than a pedagogical method that is grounded in and guided by empirical findings. Empirical research has suggested that guided discovery appears to be a better approach to promoting learning than pure discovery, and it has pointed to several approaches to promoting discovery and transfer during learning. A fruitful avenue for future research, according to Mayer, would be to explore precise mechanisms involved in discovery learning, in particular the quantity and quality of guidance that results in optimal learning performance. This type of empirical evidence generated from lab and classroom experiments can then be extended to and used to guide educational practice. Commonsense beliefs about the importance of inquisitiveness can then be separated from empirically driven educational practices, and discovery learning would then be established as a truly evidence-based educational method.

Zhe Chen and Ryan Honomichl

See also Cognitive View of Learning; Constructivism; Explicit Teaching

Further Readings

Bruner, J. (1962). *On knowing: Essays for the left hand.* Cambridge, MA: Harvard University Press.

Chen, Z., & Klahr, D. (1999). All other things being equal: Acquisition and transfer of the Control of Variables Strategy. *Child Development*, *70*, 1098–1120.

Chi, M. T. H., de Leeuw, N., Chiu, M.-H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science*, *18*, 439–477.

Dewey, J. (1938). *Experience and education.* New York: Macmillan.

Klahr, D., & Nigam, M. (2004). The equivalence of learning paths in early science instruction: Effects of direct instruction and discovery learning. *Psychological Science*, *15*, 661–667.

Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? *American Psychologist*, *59*, 14–19.

McDaniel, M. A., & Schlager, M. S. (1990). Discovery learning and transfer of problem-solving skills. *Cognition and Instruction*, *7*, 129–159.

Rittle-Johnson, B. (2006). Promoting transfer: Effects of self-explanation and direct instruction. *Child Development*, *77*, 1–15.

Siegler, R. S. (2002). Microgenetic studies of self-explanation. In N. Granott & J. Parziale (Eds.), *Microdevelopment: Transition processes in development and learning* (pp. 31–58). New York: Cambridge University Press.

Toth, E. E., Klahr, D., & Chen, Z. (2000). Bridging research and practice: A cognitively based classroom intervention for teaching experimentation skills to elementary school children. *Cognition and Instruction*, *18*, 423–459.

DISCRIMINATION

In the social sciences, *discrimination* refers to the differential treatment of people as a function of their group membership. Because many people experience

discrimination over the course of their lives, and because people may discriminate without conscious awareness, discrimination is an extremely relevant topic within educational psychology. This entry defines *discrimination* and then explains its causes and consequences. Finally, several attempts to combat discrimination will be described.

Definition of Discrimination

Discrimination includes differences in verbal and nonverbal behavior, such as when a White interviewer makes speech errors or less eye contact when interacting with an African American as opposed to a White job applicant. Discrimination may also be blatant or subtle in form, and it may be intentional or unintentional. For instance, discrimination may include obvious acts of aggression, social exclusion, differences in the allocation of valued resources (e.g., raises and promotions), or subtle acts of condescension. Discrimination may, at times, seem to be characterized by seemingly kind acts, as well as negative acts. Traditional women and elderly people, for instance, may be excessively praised for accomplishments because such accomplishments are not expected of women and elderly people. Importantly, however, such praise tends to be condescending because it has an implicit qualifier, such as "Wow, she's great, *for a woman!*"

Members of many different groups may be targets of discrimination. These groups include, but are not limited to, women, racial minorities, elderly people, gay men, lesbian women, immigrants, people who live in poverty, people with physical disabilities, and overweight people.

Causes of Discrimination

Discrimination is thought to follow from prejudice and stereotyping. Although this is often the case, stereotyping, prejudice, and discrimination are distinct constructs.

Stereotypes, prejudice, and discrimination are typically conceptualized as the three components of intergroup attitudes. Prejudice refers to the *affective* component of an attitude. It describes the way a person feels about a particular group, which can be negative or positive. Stereotyping refers to the *cognitive* component of an attitude. In other words, stereotypes are beliefs about a particular group. Like prejudice, stereotypes can have negative and/or positive content.

If someone thinks that women are warm and communal, but not influential and leader-like, then that person holds a stereotype about women. Discrimination refers to the *behavioral* component of an attitude.

Discrimination is a complex phenomenon that sometimes stems from prejudice. Some theorists, for example, argue that prejudice often is expressed through ingroup favoritism. People tend to value that which is and those who are familiar, and people tend to protect their values through ingroup favoritism. Others have linked prejudice and discrimination to personality constructs, such as the authoritarian personality (characterized by preoccupation with power, authority, and adherence to cultural ideals). Importantly, discrimination is related to the natural and adaptive mechanisms humans use to negotiate and make sense of complex social contexts. People view themselves as group members to satisfy the need to belong and to achieve positive social identities. By associating with groups and differentiating one's own group from other groups, people can belong to valued groups and achieve a positive sense of self. Ironically, adaptive self processes are often associated with discrimination, or favoring one's own group at the expense of the other groups.

There is also ample evidence that stereotypes and discrimination are linked. For instance, educational and social psychologists have conducted research on self-fulfilling prophecies, particularly in educational settings. The concept of self-fulfilling prophecy was first put forth by sociologist Robert Merton, who suggested that one person's expectations about another person could cause changes in that other person's behavior. Picking up on this idea, Robert Rosenthal and Lenore Jacobson published their famous "Pygmalion in the classroom" study. These researchers led teachers to believe that certain students in their classroom were "late bloomers," thus creating the expectancy that these students had traits that would lead them to be stars in the classroom. The so-called late bloomers were, however, randomly selected. Nevertheless, compared to a control group of students, the late bloomers showed more striking gains in IQ over their elementary school years. Rosenthal and Jacobson's seminal research inspired much theory, research, and debate. After nearly four decades of theory and research on self-fulfilling prophecy effects, and related outcomes, most acknowledge that stereotype-based expectations influence behaviors in ways that reinforce and maintain stereotypes, and the status quo more generally.

Although expectancies can be formed consciously, the self-fulfilling prophecy effect can be considered unconscious in that people don't purposely set out to change the behavior of others through their expectations. However, stereotyping can also influence discrimination at an unconscious level. For instance, stereotypes can be automatic in that when a person sees someone else, he or she automatically processes certain key information about that person (the person's age, race, and gender) without conscious awareness. This categorization activates stereotypes associated with particular groups. For example, when someone categorizes another person as African American, stereotypes about African Americans become activated. Thus, the person who is doing the categorizing may behave in discriminatory ways based on the group to which the target was ascribed. For instance, the perceiver may stand further away from the person being categorized than if that person had been categorized as White.

Consequences of Discrimination

Discrimination often has damaging consequences. For instance, many scholars have theorized that discrimination leads to decreased confidence and self-esteem. Although this idea has been challenged, evidence remains that discrimination leads to harmful emotional effects. These include anxiety, self-doubt, decreased confidence, increased anger, and fear of confirming negative stereotypes.

In addition to psychological consequences, discrimination may result in tangible, everyday harms. One of the most frequently documented cases of discrimination involves "equal pay for equal work." Although women frequently do not realize that they are the victims of discrimination because they do not have access to the financial records of men in their companies and therefore do not know that they are earning comparatively less, they still earn an average of only 76 cents for each dollar earned by men.

Discrimination can also adversely affect academic performance. Members of stigmatized groups often underachieve in academic settings. This could be due to many factors, including the students' fears of confirming stereotypes about their groups. This phenomenon is known as *stereotype threat.* Thus, stereotyping leads to decreased academic performance, which in turn leads to discrimination, as grades and standardized test scores influence hiring and admission decisions. This results in a system whereby members of stigmatized groups experience discrimination due to their underperformance on academic tasks, which is at least partially the result of stereotypes about their group.

Combating Discrimination

To reduce discrimination, the U.S. government has implemented several policies, including affirmative action. Affirmative action began in 1965, when President Lyndon B. Johnson signed a mandate requiring federal contractors to not "discriminate against any employee or applicant for employment because of race, color, religion, sex, or national origin" and to "take affirmative action to ensure that applicants are employed, and that employees are treated during employment, without regard to their race, color, religion, sex or national origin."

Affirmative action has created controversy with regard to employment and admission to institutions of higher education. In the 1978 *Regents of the University of California v. Bakke* case, the Supreme Court ruled (five to four) that universities could consider racial heritage when making admissions decisions. In 2003, the Court upheld this decision in cases involving the University of Michigan by ruling that race can be one of many factors considered in admissions. The Supreme Court further ruled, however, that point systems such as those used by the University of Michigan's undergraduate admissions program had to be modified because they do not provide individualized consideration of the applicants.

Nongovernmental intervention programs have also been implemented across the United States. For instance, one study found that African Americans performed significantly better in school when asked to write about values that were important to them as a way to reaffirm their sense of personal adequacy. Such intervention programs suggest that the harmful effects of discrimination can be reduced through strategies targeting stigmatized group members.

Larisa Heiphetz and Theresa K. Vescio

See also Cultural Diversity; Ethnicity and Race; Gender Bias; Stereotypes

Further Readings

Allport, G. (1954). *The nature of prejudice.* Oxford, UK: Addison-Wesley.

Fiske, S. T. (1998). Stereotyping, prejudice, and discrimination. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 2, pp. 357–411). New York: McGraw-Hill.

Jussim, L., & Harber, K. D. (2005). Teacher expectations and self-fulfilling prophecies: Knowns and unknowns, resolved and unresolved controversies. *Personality and Social Psychology Review*, *9*, 131–155.

Merton, R. K. (1948). The self-fulfilling prophecy. *Antioch Review*, *8*, 193–210.

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*, 797–811.

Tajfel, H., & Turner, J. C. (1986). The social identity theory of inter-group behavior. In S. Worchel & L. W. Austin (Eds.), *Psychology of intergroup relations.* Chicago: Nelson-Hall.

U.S. Department of Labor. (2006). Executive Order 11246, as amended. Retrieved October 30, 2006, from http://www.dol.gov/esa/regs/statutes/ofccp/e011246.htm

DISTANCE LEARNING

Distance learning is a term in wide use today. Educators also refer to it as *distance education* and, in some settings, *distributed education.* For the purposes of this entry, *distance learning* is defined as the communication over distance between teacher and student mediated by print or some form of technology designed to bridge the separation between teacher and student in space or time. Advances in information and communication technology are changing the manner in which instructors have traditionally conducted distance learning, and these changes are providing many transformational possibilities for all levels of education. With the development of many online tools and the easing of prices for handheld computers and audio/video players, students are increasingly able to shift their distance learning experience not only beyond temporal necessities, but also into new physical environments. Research libraries with access to full-text documents are as available to students as are lectures and symposia either streamed live or delivered asynchronously. Educators could argue the case that advances in information and communication technologies may make distance learning even more interactive than face-to-face teaching and far less distant than once considered.

After an examination of the origins of distance learning and a review of its basic features, this entry examines some of the effects on pedagogy of course management systems and other telecommunications tools that are transforming the nature of schooling, lifelong learning, and communities of learners.

Origins and Evolution of Distance Learning

Distance learning has been available in one form or another for hundreds of years. One of the earliest examples of distance learning occurred in England in the 1840s. The Pitman Company offered training in shorthand through a series of lessons mailed to students across the country. In hindsight, this was very much a one-way, noninteractive approach to distance learning.

Until recently, public interest in distance learning was especially high only where there was a widely distributed student population. One of the more famous modern examples of distance learning took place in Australia. Beginning in 1951, the School of the Air officially opened to broadcast, by radio, lessons to the children of the Outback. Beginning with one-way transmissions, coordinators soon added a question period to follow the broadcasts. Interactivity, even in the beginning stages of technology-enhanced distance learning, was highly valued.

Today, many institutions of higher education, both public and private, are making use of distance learning to broaden the reach and scope of their particular curricula. Of course, it is very important that students who participate in distance learning are self-motivated and able to work independently, but teachers also have a particularly vital role to play in the process of distance learning and its overall evolution.

Basic Features of Distance Learning

Distance learning and distance education programs are experiencing a boom of sorts with the advent of the Internet and the ability to transmit increasingly large audio and video files over increasingly available and accessible bandwidth. A proper examination of distance learning must begin by determining how distance learning differs from face-to-face learning. In fact, many research studies focusing on the efficacy of distance learning compare it with face-to-face learning. For that reason, researchers view many of the tools and affordances of modern distance learning

as supplements to or extensions of traditional or face-to-face learning experiences.

One of the most basic features of traditional or face-to-face learning for the student is the ability to listen to and interact with the instructor. Distance learning, using a variety of tools and technological affordances, can replicate the primary experience of listening to an instructor using audio files downloaded and played on a computer or portable listening device. Advances to this technology also enable students to view the instructor and available visual aids on convenient handheld devices. The availability of such files leads to another feature of distance learning, the ability to shift the learning experience out of classroom into a different time and place. This distribution of the learning experience is known variously as "phase shifted learning," "asynchronous learning," and "any place any time learning."

Transactional Contact

Another important and essential element easily accomplished in face-to-face instruction is transactional contact, or the ability of the instructor to interact with the student. With the advent of several software tools, instructors can shift their transactional contact with students beyond the classroom. Instructors accomplish this in two different ways.

In one way, the instructors retain the ability to speak and even see the student over distance in real time. Instructors accomplish this interactivity relatively inexpensively with a simple web camera working in concert with instant messaging software. There is available to instructors a range of sophisticated video conference systems that transmit video and audio over the Internet. These systems make use of an Internet networking protocol widely referred to as "Video over IP." An earlier version of this protocol, which is still in use, makes use of ISDN or Internet Subscriber Digital Networks to allow for two-way video conferencing.

The other way of maintaining close transactional contact between the instructor and the student is asynchronously using software-driven quizzes and reviews of instructional material that provide automated feedback to the student. Instructors can provide this feedback in varying intervals, for example, after the submission of responses to a series of questions or upon the submission of a response to a single question. Variations of this technique are gaining greater

use within distance learning environments. One example is to provide the conditional release of assignments so that before students can progress to the next stage or segmented element of learning, they must demonstrate, in some manner, a degree of understanding the material. The ability to automate this mastery learning technique is a tool that is causing some instructors to revamp the pedagogical structure of their courses.

Distributed Learning Materials

One of the basic elements of distance learning that technology has enhanced recently is the ability for instructors to distribute reading materials for courses quickly and efficiently as electronic documents accessible across computing platforms. Many instructors use the ubiquitous portable document format also known as PDF, which requires a free software reader. The instructor creates these documents, gathers them from online sources, or scans them directly from a primary source.

It is not only documents that can be distributed in this manner, but also spreadsheet files, still photographic images, software programs, and even movies. Such a wealth of resources, commercially available and user created, allows for the possibility of increased opportunities for reinforcement or for the reteaching of concepts. Resources such as these are known as Learning Objects. Many distance learning projects may store them in what is known as a Learning Object Repository (LOR) and make them available to other institutions and professors so they do not need to reproduce work that another colleague has already done.

Modification of Pedagogy

Such a wide availability of tools and resources also allows the instructor to guide the student to the pedagogical experience and learning environment that is most suited to the learning style of the student. Such an explicit design of the learning experience, one that the instructor can tailor directly to the learning style of the student, is a great challenge to instructional design experts and still the goal of many creative instructors.

In addition to modifying instruction to learning styles, the latest technological affordances improve instructors' ability to present materials to accommodate physical impairments or disabilities. For example,

students may use text-reading software to read and review class materials, and students with visual impairments can use adaptive technologies to increase or improve the visual presentation of materials using the adaptive technologies of their choice.

Course Management Systems

With so many tools available to instructors, an essential element of distance learning has become the Course Management System, sometimes known as CMS, or a Learning Management System, or LMS. These management systems make available to instructors a variety of tools they may opt to use to make the distance learning experience uniquely tailored to the course content. The potential tools and affordances, of course, can be broken down along various continua, such as content rich versus content poor or synchronous versus asynchronous.

Discussion Lists

Discussion lists are primarily asynchronous, text-based opportunities for students to respond to questions and discussion topics and to interact with each other in the familiar environment of their e-mail program of choice. Innovations in this area are allowing students to leave video or audio messages for each other and their instructor.

Threaded Discussions

Commonly known as discussion boards, threaded discussions are web based, and the instructor can set the degree to which students may respond by limiting length of response or the number of responses. One benefit of the threaded discussion is that instructors can populate the discussions with small groups of students and change those groupings as needed.

Chat

Chat, sometimes known as the Internet Relay Chat (IRC), allows multiple participants in the course to enter a text-based environment to correspond with each other and the instructor. An unfortunate drawback of this tool is that students sometimes have difficulty following multiple conversations simultaneously. For this reason, many course management systems allow students to color code their text to make it more easily recognizable by others in the conversation. Some course management systems allow the use of images or avatars to represent the speaker in a conversation.

Internal Communication

Course Management Systems sometimes employ their own e-mail services that are accessible only when the students log on to the course. Also widely available are tools for instant messaging or paging another individual in the class when students log on to the system. Recent iterations of this tool allow the posting of video messages.

Presentation Tools

Instructors who employ visual slide presentations in their face-to-face teaching can post the presentation in its entirety to the online course. They can also use many available tools to capture their speech and include their narration with the presentation or post their narration as an enhanced podcast that students may review asynchronously. Using other widely available tools, instructors may also capture a sequence of keystrokes to illustrate a technique or the steps necessary to operate a software program. They may also use an electronic white board enabling the instructor to interact with graphical elements in the course using a digital marker.

Digital Content

Distance learning also enables students to participate in the creation of learning objects that the instructor can share with the class or other colleagues. One of these techniques is to create a repository of digital objects for analysis and critique by other students in the course. Depending upon the nature of the digital objects, such as images, movies, software, and Web pages, there are several systems by which students may access these objects. For example, students may create a glossary for use by fellow students. Students may also make use of a tool known as a wiki, which will allow classmates the opportunity to edit a living document as the knowledge base is refined, improved upon, or corrected.

Instructors also may make use of the repository of digital teaching tools created by their peers in the field. Instructors add these digital tools to the digital repository and tag them with key words, sometimes known as metatags, which allow the repository to store and retrieve the objects with ease. The rise of

this method of exchanging digital teaching tools has led to the development of standards for the industry. The Sharable Content Object Reference Model (SCORM) is the standard for tracking records in learning management systems. The benefit of having this sort of repository is that the users, both students and instructors, can go to the database, or repository, to view or download tools.

Issues Related to Distance Learning

In some institutes of higher education, much has been made of the increasing absenteeism of students because of professors and instructors being able to distribute their lectures in digital format. Some instructors see this loss of students in a physical building as anathema to their understanding and perception of themselves as teachers. With so much of the face-to-face classroom experience enhanced by myriad ways of gaining feedback from the class, from laughter to subtle signs of discomfort, the instructor loses the opportunity, in some respects, to modify the presentation of concepts to his or her students. For example, an instructor who is speaking about a social networking site on the Internet as a tool for research may have his or her lecture interrupted by a comment from a student who may be aware of a similar site that the professor could then include in future lectures or use to forward his or her own research agenda.

Assessment and Shifting Strategies

One of the most common features of any sort of educational interaction is the ability of the instructor to perform either summative or formative evaluation to determine whether students have understood the concepts of the course. In a face-to-face teaching environment, instructors conduct performance evaluations informally or through assigned projects, whereas summative evaluations are usually the domain of the pencil-and-paper test for examination.

In the online distance learning environment, one of the earliest implementation issues to arise reflected the difficulties of assessing work done by students beyond the scrutiny and gaze of the instructor. Responses to the difficulty of determining whether the students' work is their own have gone in two directions. One way of reducing the possibility that students might share answers is for instructors to provide limited timed testing situations whereby students have only one opportunity to answer questions.

Such tools and techniques to restrict the possibility that students will do summative evaluations in concert with their classmates have their place. However, other instructors have taken a different and potentially transformative approach in response to this problem. That approach has been to change the nature of the assignments students are doing in class. The changes have used many of the tools and affordances of the distance learning environment and course management system employed. Responses to the issues related to assessing learning have led to a major shift in pedagogical models.

Collaborative Teams

One of these transformational changes in pedagogy has been the development of collaborative teams in which students work on a project together online and evaluate each other's contributions to the sum total of the project. The overarching concept behind this approach is that in order to accomplish the project, as a team, members needed a full conceptual grasp of the core concepts. Students would then demonstrate this understanding of the concepts through methods incorporating higher-level thinking skills. Simple restatement of basic facts, which may have been a standard feature in a face-to-face classroom, is a technique that instructors are replacing with project-based activities involving synthesis of new material, evaluation of the work of others, and activities typically associated with complicated thinking tasks.

The Role of Community in Distance Learning

In distance learning, two opposing pedagogical perspectives have had an impact on instructional design: symbol processing and situated cognition. With symbol processing, the teacher transforms abstract ideas into concrete images that he or she presents to the learner via some sort of medium such as television, radio, or the Internet. The job of the learner is to perceive, decode, and store the information from the concrete images.

With situated cognition, the student interacts with both the problem and the construction of a solution to the problem. Students engage in a process of inquiry and draw conclusions from that process. Instructional

designers develop presentations that enable learners to construct new knowledge by interacting with the instruction.

In some respects, distance learning recreates the teacher–student interaction of a face-to-face classroom, but it does so through either synchronous or asynchronous means. Although generally considered primarily asynchronous with the occasional synchronous interactions through a variety of technological means, recent studies have suggested that one of the better models for distance education includes some face-to-face teaching and direct interaction with students. Instructors can effectively place this face-to-face interaction at the beginning of the course or spread it throughout. The physical interaction with students allows the online instructor an opportunity for rich feedback and encourages a more robust learning community.

Virtual learning communities, facilitated by the instructor, increase interactivity between teacher and student and encourage students to interact with each other and with the course content in the online environment. In the case where many thousands of miles separate students, instructors can use inexpensive technological tools to encourage the sense of a learning community.

Lifelong Learning and Distance Education

Although the traditional perception of distance learners is that of the university student or adults engaged in lifelong learning, there are many interesting forms of distance education occurring at the elementary and middle school levels of education. Mostly these are in the form of enrichment activities or telecommunication projects involving the entire class. There are many online sources for projects such as these if teachers were to express an interest in participating. Some projects make use of televised instructional sources, whereas others are entirely text based through electronic mail or through the Internet. Recently, students at this stage in their education have participated as data gatherers in a larger network of students, sometimes global in nature, engaged in discovery projects.

At the secondary school level, rural classrooms and small school districts are making use of distance education courses to help students take advanced placement courses that their school district is not able to offer because of limited enrollment or funding. Educators can also offer these opportunities for taking courses outside the core curriculum in areas such as vocational education or foreign language to students who are not able to attend school because of a disability or because they have chosen to be homeschooled.

One concern that some educators have recently raised is that the boom in distance education may disenfranchise students lacking the skills of self-motivation and discipline. The availability of technological tools enabling the growth of distance learning does not ensure that instructors are well trained either in the development of online courses or in their delivery. Although a high level of interactivity with students can overcome deficiencies in some online courses, some states have considered mandating that at least one online course be required for graduation. In this way, students will gain familiarity with distance education methods and delivery systems, a method of teaching they will undoubtedly encounter in their lives beyond high school.

Digital Literacy

Although television and the Internet can motivate students and stimulate an interest in learning, exciting visuals may serve to focus students' attention more on the form than on the message and underlying meaning. This concern has led to much discussion about the necessity of developing a new kind of literacy in which students are able to distinguish between fact and persuasion, between valid information and propaganda, and between reliable sources and biased opinion. Educators have begun calling this phenomenon *digital literacy* or *Internet literacy.*

Advantages and Disadvantages

There appear to be as many advantages to distance learning as there are disadvantages for each of the stakeholders involved. For students, distance learning renders distance from training centers less relevant. Distance learning has flexibility, accessibility, the ability to create a self-paced learning environment, and less wasted time. That said, there is a loss of direct interaction with the instructor and classmates, a potential loss of immediate feedback, the risk of a higher rate of failure, no access to a physical library, and the potential limitations of the student's own computer system.

For the instructors, there is the possibility of a larger audience, less classroom repetition, more time available to upgrade the course, and the acquisition of

a new teaching experience. However, distance learning changes the dynamics of the classroom environment, there is no visual contact with students, it is sometimes difficult to evaluate students' work, and there is a huge workload at the beginning of the course.

For the institution, there is no need for a physical building, there is potential for international clientele, and the possibility of additional income streams. However, the institution may be in a position to have to reevaluate how it grants credits, it may have to redefine the role of the professor's duties, and it may have to define the rules and procedures for online courses. Additionally, the institution must police an increasingly complex environment regarding copyright and digital rights management.

Distance Learning 2.0

Distance learning is turning a huge corner as scholars and nonscholars alike are using newly emerging social networking tools. In the most recent versions of distance learning, instruction held within classrooms is gradually shifting toward classes conducted virtually, within a computer-mediated communication environment. What is happening now is a steady shift toward learning beyond the classroom in ways that we might excuse the most forward of thinkers for not envisioning.

Sometimes referred to as *e-Learning 2.0* or the *read/write web*, with social networking tools requiring almost no specialized software such as blogs, wikis, and podcasting, the nature of distance learning is transforming. It is moving from the formal model of the instructor in charge of a curriculum to that of an informal gathering of like-minded researchers, both professional and amateur, communicating with each other through direct dialogue or through a trail of references. Using any of a large range of tools, students, classes, or even whole departments can add tags to articles found online and store these tags as links to the articles in such a way that they can share their finds with their larger community of scholars.

The use of interactive blogs creates a social network that moves beyond the very concept of what a course is and can move learning to a place beyond the classroom and even beyond the structured curriculum. Even within a structured curriculum, online blogging tools, with their capability of allowing comments, will enable outside experts to add comments and insights to transform and increase the potential

for enrichment to the distance learning environment. Such tools support increased feedback from instructors, but there is even greater potential for the development of community-centered instruction. Such tools move education from theoretical notions about social cognition and the development of communities of practice to concrete experiences for teachers to create online learning communities.

Of course, the educational interactions do not simply occur without planning on the part of the instructor. Despite the increased participation by all but the most reluctant students to participate in class discussions in an online setting, there is still a need to set guidelines and provide educational leadership and facilitation.

The tool known as a *wiki* enables many users to add to a base of knowledge drawn from their own research and investigations into a topic. Instructors can add a line or two of information with a link to an outside reference source that students can then modify and inform with many other additional sources. This knowledge construction is more actively engaging than other, more passive forms of pedagogy. With the advent of new online tools, such wikis are becoming commonplace classroom tools in the distance learning environment.

Michael H. McVey

See also Learning Communities; Lifelong Learning; Media Literacy

Further Readings

Bawden, D. (2001). Information and digital literacies: A review of concepts. *Journal of Documentation*, 57(2), 18–59.

Brown, R. E. (2001). The process of community-building in distance learning classes. *Journal of Asynchronous Learning Networks*, 5(2), 18–35.

Dede, C. (1996). The evolution of distance education: Emerging technologies and distributed learning. *American Journal of Distance Education*, 10(2), 4–36.

Harasim, L., Hiltz, S. R., Teles, L., & Turoff, M. (1996). *Learning networks*. Cambridge: MIT Press.

Kellner, D. (2000). *New technologies/new literacies: Reconstructing education for the new millennium.* Retrieved October 4, 2002, from http:// www.gseis.ucla.edu/courses/ed253a/kellner/ newmedia.html

Reigeluth, C. M., & Duffy, F. M. (2007). Trends and issues in P–12 educational change. In R. A. Reiser & J. V. Dempsey (Eds.), *Trends and issues in instructional*

design and technology (2nd ed.). Upper Saddle River, NJ: Pearson Education.

Rovai, A. (2002). Building sense of community at a distance. *International Review of Research in Open and Distance Learning*, *3*(1). Retrieved August 15, 2007, from http://www.irrodl.org/index.php/irrodl/article/view/79/152

Divergent Thinking

Creative thinking is often dichotomized into thinking that is convergent and thinking that is divergent. Of the two, convergent thinking is much simpler to measure and operationalize, whereas the latter, a cognitive process that engenders a variety of novel and unconventional products, has proven far more difficult to model, measure, and predict. To address the complexity of this difficulty, it may be helpful to clarify what divergent thinking is not. By framing a discussion of divergent thinking upon a theoretical delimitation of convergent thinking, one may develop a firmer foundation for a description of cognitive processes that are, by negative comparison, divergent.

The objective of convergent thinking is to generate a conventional, consensually agreed-upon solution to a problem. The problem presents as an initial triggering mechanism or process input—be it a test prompt, an assignment question, or an event—that generates the production of a correct or an incorrect problem solution. Convergent cognitions, therefore, are necessarily goal-directed, and thinking tends toward a linear, serialized sequence of subprocesses channeled via feedback mechanisms such as task monitoring, evaluation, and output verification. Traditional models of the creative process, such as Graham Wallas's *Preparation → Incubation → Illumination → Verification* and John Dewey's *Difficulty Felt → Difficulty Defined → Information Surveyed → Solutions Suggested → Outcome Considered* reflect a deep-structure supposition that creative thinking is a linear, goal-directed process of solving a problem. In addition, to the extent that the goal of convergent thinking is to achieve objectives external to the thinker—that is, success and failure are not based on internal, personal criteria, but rather success is based on the opinions, judgments, and preexistent correct solutions of external agencies—affective engagement with the problem to be solved may, at least initially, tend toward an extrinsic rather than an intrinsic basis for individual motivation and problem relevance.

In contrast to convergent thinking, divergent thinking does not rely on a problem input to initiate cognitive process. A sculptor may "seek" a form in the grain of a block of marble or a poet may "chase" the meaning of a memory triggered by a chance image that then initiates clusters of associated images, memories, and word sounds. These associative clusters (conscious and otherwise), domain-specific facts, procedures and concepts, chance discoveries, environmental intrusions, and dynamical process changes need not work in a serialized fashion, but rather, subsystems may parallel process in simultaneity. As a result, in order to be successful, the divergent thinker has to remain sensitive to competing, often chaotic interactions while maintaining the cognitive equilibrium necessary to generate a coherent artifact.

The above description suggests that divergent thinking processes engender an artifact that is, in a sense, an emergent by-product of thinking, not the end product of problem solving. Indeed, this product as by-product paradigm implies that divergent production, unlike convergent production, renders impossible the prediction of a final outcome or the prospect of recreating cognitive and affective processes back to an initial impulse. Another way to consider this phenomenon is, as the divergent thinker creates, that which she or he creates affects and changes the creator, which in turn affects and changes that which is created. A definition of divergent thinking, therefore, must take into consideration the following subconstruct: Divergent thinking is simultaneously and reciprocally a creating process and a learning process. Finally, apropos to the affective domain, if divergent thinking has the capacity to be internally triggered, and ideas are accepted, rejected, and integrated as guided by personal, internal, intuitive proclivities, then process-associated intrinsic motivation may play a greater role in divergent thinking than it could, comparatively, in convergent thinking.

To summarize, divergent thinking is a complex, associative process of bringing alternative, novel, and unconventional ideas to emergence through activating and accessing both conscious and subconscious subsystems and processes. Herein, the production as by-product construct is characterized by two attributes: First, the act of creation imparts learning-induced change upon the creator, and second, divergent thinkers use an open-ended, often chaotic process of seeking and discovery in contrast to a closed-ended, consciously task-monitored, linear process of solving and answering.

Putting aside recognized issues of discriminant validity, construct validity, and interjudge reliability (e.g., how precisely may one determine what is or is not conventional?), that traditional psychometric tests artificially initiate the cognitive process to be measured via problem prompts suggests one hypothesis as to the limited predictive ability of these tests: External prompts may have the effect of negatively influencing the affective domain of creativity and directly inhibiting the very divergent, non-goal-directed seeking processes that many measures of creativity attempt to capture.

William T. Akers

See also Creativity; Intrinsic Versus Extrinsic Motivation

Further Readings

Dewey, J. (1910). *How we think.* Boston: Heath.

Glover, J. A., Ronning, R. R., & Reynolds, C. R. (Eds.). (1989). *Handbook of creativity.* New York: Plenum.

Guilford, J. P., & Hoepfner, R. (1971). *The analysis of intelligence.* New York: McGraw-Hill.

Kim, K. H. (2006). Can we trust creativity test? A review of the Torrance Test of Creative Thinking (TTCT). *Creative Research Journal, 18*(1), 3–14.

Lubart, T. L. (2001). Models of the creative process: Past, present and future. *Journal of Creative Behavior, 13*(3), 295–308.

Runco, M. A. (Ed.). (1991). *Divergent thinking.* Norwood, NJ: Ablex.

DIVERSITY

In the field of educational psychology, *diversity* primarily refers to differences across individuals and groups. In order to talk about differences, there is often an assumption about what is normal, typical, or mainstream. Educational psychologists have been interested in understanding human life by identifying universal tendencies. Diversity is important to this task because it is the concept through which hypothesized human universals might be better understood. For example, an interest in diversity helps psychologists understand why there is patterned variance across individuals and groups given particular hypothesized explanations for cognitive development. Steele's important work on stereotype threat has explored how performance (including cognitive tasks) can be influenced by social stereotypes. Because

social stereotypes work off of, and reinforce, social expectations for groups of people (by race and gender, for example), diversity is an important theme.

In mainstream educational psychology, it is rare to find the word *diversity* in the index or in a chapter title. Its position in educational psychology is somewhat indicated by that fact. It is, nonetheless, not a totally invisible concept. Diversity has found its way into educational psychology because researchers are interested in the ways in which society and social forces influence the individual. As such, racial and gender diversity as social forces, for example, have gained the interest of educational psychologists. Moreover, concern for social justice and attempts to understand social phenomena such as violence and prejudice have led educational psychologists to look more closely at social diversity. There are four trajectories in educational psychology in which concepts of diversity have become of vital interest. Each of the five trajectories, described below, implies slightly different emphases in the definition of *diversity*. They are presented below following a rough historical chronology, but the reader should realize that these conceptions of diversity have influenced and affected each other and can all be seen in contemporary educational psychology.

In the first place, psychologists are interested in developing theories and research explaining and describing human universals. Although educational psychologists tend to be attentive to individual differences and deviations, the idea of "human" was not socially complicated. There were many assumptions of universality left unquestioned in the research designs and theories. For example, Kohlberg initially studied "human" moral development by examining only men, assuming that men and women were the same. Thus, diversity enters the scene of educational psychology as a way of identifying greater complexity in human phenomena by questioning the assumptions of universality. Researchers' capacities to understand human universals and make claims that best represent the population to which they are generalizing are strengthened when those researchers question the universal assumptions at the outset so that those assumptions are not blindly built into research designs and theories.

This change in the field led to including diversity as a potential feature of the research design in educational psychological studies. In this second trajectory, *diversity* is a broad term covering a range of discrete variables. These variables generally reflect the attitude,

scope, and place of diversity within the field of educational psychology. The variables in educational psychology that are most often linked to the concept "diversity" include, for example, race, gender, culture, and sexuality. Categories of diversity entered the research of educational psychologists in this way. A person's social experiences of race, gender, culture, and so on have been found to affect personal psychological phenomena. For example, Clark and Hatfield have reported differences in receptivity to sexual offers among college-age students based on gender. Studies will now typically report the racial and gender characteristics of their subjects even if their findings are meant to generalize more broadly and even if these social categories are not considered variables in the study. This allows the readers of the research to consider the role of diversity in the research design. The term *diversity*, here, refers to a set of social categories within which one can subdivide people into easily identifiable groups. This has been a very important trajectory in educational psychology because it allows us to explain patterns of difference across humans in order to better understand potential human universals and their deviations as they are relevant for education. The tendency in educational psychology is to explain deviations from human universals as a function of individual differences. As a result of this trajectory, educational psychologists developed interests in topics such as gender identity development, homosexual identity development, and the effects of poverty on development.

Third, psychological concepts of the "individual" became more complicated in educational psychology through the notion of group membership. Relevant categories of diversity involved distinguishing groups from one another. This became an important departure in educational psychology, where developmental models of human characteristics have had a broad and enduring impact on psychologists' conceptions of humans. With developmental models, age and stage constitute primary categories of internal difference, but these are not conceptualized as characteristics of one's membership in a social group. Tajfel and Turner began studying something they referred to as social identity. *Social identity* is a composite of three aspects—social group categorization (including self-categorization), self-identification with a group, and social comparison with other groups. Social identity theory and findings suggest that groups are not just something externally constructed around individuals, but instead reflect something inherently constitutive of individuals' claims

about themselves. In other words, Tajfel and Turner piqued the interest of educational psychologists in investigating the possibility that individual differences did not necessarily precede group memberships and affiliations. When seen as variables, the characteristics of diversity are theorized as external to human universals and individual characteristics, but nonetheless descriptive and interesting for understanding psychological phenomena, such as responses to attractiveness. With this second trajectory, diversity across groups is fundamental to both the sense of group identity and the sense of uniqueness that individuals will claim about themselves. The term *diversity* here contrasts with sameness. This trajectory has led to a lot of research on ethnic identity, ingroup/outgroup phenomena, and psychology of gender, and it has likewise influenced the ways educational psychologists include diversity in their research and theorizing.

Fourth, the concepts of diversity are relevant to the research on intergroup relations, including ingroup/outgroup behavior, scapegoating, stereotype threat, prejudice, and other theoretical/empirical phenomena. In these cases, diversity reflects the notion of difference between groups as well as attitudes and responses to those differences, including attitudes and responses that are influenced by majority/minority relations. Categories of diversity are not mere variables for this body of work, but rather diversity itself is at the heart of the social phenomena of interest. There is no way to understand the basic concepts of intergroup relations, including the above list of phenomena, without the concept of diversity as indicative of group difference. For this work, diversity is most often conceptualized as interindividual differences.

Following World War II, during the rise of social psychology in the United States, studies on conformity and authority were complemented by studies on ingroup/outgroup phenomena and scapegoating. These latter interests sought explanations for the Holocaust through ideas that depended on intergroup and interindividual differences. Studies that indicated the force of arbitrary group affiliation on ingroup/outgroup behavior and preferences led researchers to investigate the not-so-arbitrary group affiliations at work in people's everyday social lives. Much current research and related theoretical ideas have emerged from this trajectory. For example, Steele launched a very important set of contested studies on stereotype threat. These studies examine the extent to which a person will underperform when his or her social group is negatively

stereotyped. The term *diversity* here refers to differences across groups and the social ramifications of those differences when social power is also evidenced. Aronson developed a popular educational tool for mediating the negative consequences of group differences and magnifying the social value of those differences—he referred to this as the Jigsaw. With both Steele's and Aronson's work, social psychology has become engaged in the critical efforts to address the negative social impacts that inequality across diverse social groups has had on society as a whole and most particularly on individual members of society.

Finally, in contrast to the move toward interindividual differences, there is a postmodern approach to educational psychology reflected in the work of Gergen. This trajectory appropriates the ideas of diversity to explain that intraindividual phenomena are influenced by diverse social forces. Gergen takes social identity theory in a different direction by arguing that any one given individual identifying with any number of social groups—with equally many stereotypes, perspectives on truth and experience, and so on—becomes "saturated" with diversity from the inside and the outside. Moreover, Gergen specifically attributes the patterned changes in intraindividual psychology to the increasingly complex and diverse social world. Gergen's approach is differently empirical. Gergen acknowledges forms of evidence not traditionally accepted within the field of social psychology; these new forms of evidence include such cultural products as films and artifacts. Gergen argues against the idea that people's minds represent or reflect the external world. The diversity of the external world does not solely influence individuals nor vice versa, but instead, the internal and external worlds are integrated in complex and contested ways. The image such ideas produce is more akin to a collage of interindividual and intraindividual diversity rather than a dichotomy. The emphasis on diversity as "perspectives" is most relevant for postmodern and poststructuralist social psychological work. From this way of thinking, the term *diversity* here refers to the multiplicity of truth claims authentically and sincerely asserted from across and within individuals as persons and as members of complex social groupings.

These more recent trends in the way diversity is conceptualized in educational psychology reflect a more critical tendency in educational psychology than was evidenced in its early days. Empirical findings and theoretical developments are linked to efforts to answer social problems so as to improve the psychological life conditions. The five trajectories in which diversity is found relevant to educational psychology also then reflect a history of educational psychology. Taken all together, we see diversity largely referring to group-related differences that have both inter- and intraindividual points of interest. Across the trajectories, there is a broadening of how diversity is conceptualized, moving from the idea that diversity was the sum total of individual differences across human universals toward the idea that it was a social phenomenon that transcended individual difference while also involving it.

Barbara Korth

See also Cultural Deficit Model; Cultural Diversity; Culture; Ethnicity and Race

Further Readings

Aronson, E. (2001). *Nobody left to hate.* New York: Owl Books.

Clark, R. D., & Hatfield, E. (1989). Gender differences in receptivity to sexual offers. *Journal of Psychology and Human Sexuality, 2*, 39–55.

Gergen, K. (2000). *The saturated self: Dilemmas of identity in contemporary life* (Reprint ed.). New York: Basic Books.

Gilligan, C. (1977). In a different voice: Women's conceptions of self and morality. *Harvard Educational Review, 47*, 481–517.

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental and Social Psychology, 35*, 4–28.

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist, 52*, 613–629.

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69*, 797–811.

Tajfel, H., & Turner, J. (1986). The social identity theory of inter-group behavior. In S. Worchel & L. W. Austin (Eds.), *Psychology of intergroup relations* (pp. 7–24). Chicago: Nelson-Hall.

DIVORCE

Divorce is defined as the legal dissolution of a marriage. As a legal action, divorce culminates in a final

divorce decree, which outlines for each individual his or her obligations for division of property and assets; support, if indicated; and provisions for children. Separate from legal implications, divorce is a transition that implies logistic and psychological challenges for all involved. Common challenges for adults going through the divorce process are processing grief, restoring self-esteem, managing anger, and easing emotional pain for their children. Common tasks for parents going through the divorce process are developing an identity as a divorced parent, becoming an effective single parent, dealing with their child(ren)'s anger, and establishing a workable co-parenting relationship with their ex-spouse.

Divorce is important to educational psychology because it is associated with family and child factors that have an indirect effect on children's educational outcomes. Thus, this entry focuses on divorce within the context of families with young children. This review first provides a brief, recent historical overview of divorce in the United States. This is followed by highlights of typical child divorce effects, as well as developmentally specific child divorce effects. Finally, the impact of divorce on the learning environment is presented, with a focus on school-based programming to support children and families.

Historical Overview

The divorce rate rose from 0.3 per 1,000 population in 1967 to its highest rate of 5.3 per 1,000 population in 1979. The change in social attitudes related to modern urbanization, including increased education and employment for women combined with smaller numbers of children per household, and the change in legal guidelines that allowed for no-fault divorce primarily account for this considerable change.

Currently, 51% of all U.S. marriages end in divorce. Often, children are involved, and more and more families with the youngest children are finding themselves in the predicament of divorce. For children who experience the divorce of their parents before age 12, at least two thirds experience it by the time they are 6 years old. Of all children who experience the divorce of their parents before the age of 6, 40% do so in the first year of life, 20% between 1 and 2 years, 15% between 2 and 3 years, 10% between 3 and 4 years, and 15% between 4 and 5 years. Thus, divorce is a common occurrence in the lives of infants, preschool, and early school-age children.

Specific to the context of divorcing families with young children, the judicial system has had the task of considering divorce in the context of children's best interests. This focus dates back to the 19th century, when it was initially legally specified that children have rights independent of their parents. This assumption was the basis for the later "best interests of the child" principle. The term was coined in the 1925 *Finlay v. Finlay* case and places judicial officers in a parental role to the family, determining what scenario best suits the needs of a given minor child. During the early 1900s, the realm of parenting responsibility was predominantly determined by gender, with mothers being responsible for nurturance and fathers being responsible for discipline. The general consensus in a divorce proceeding resided in a mother's parenting role carrying greater weight for children's development than the father's. As a result, the majority of residence decisions placed children with their mothers. This earlier approach was mirrored in the 1970s with the Uniform Marriage and Divorce Act, which emphasized the emotional needs of children in the "tender years" as being best served by residing with their mothers. This presumption has been scrutinized for its unconstitutionality and has gradually been replaced by a case-by-case review among judicial officers applying the best interests of the child principle.

Currently, the principle primarily pertains to decisions about child residence and parental access. These decisions are established in the form of parenting plans with outlined parenting time responsibility. What remains the same is the aim of the principle, which is to reduce the concept of parental ownership of children and minimize a competitive stance in litigation, with resulting judicial decisions articulated based on children's needs rather than parents' rights. The basis for understanding what is in children's best interests has evolved out of collaboration between the legal and mental health professions. Specifically, between 1973 and 1986, and revised in 1996, the principle encompassed the idea that custody decisions should both safeguard the child's need for continuity of relationships while reflecting the child's sense of time. Some states used a criterion set forth in Michigan as a check against this conceptual definition of the best interests of the child principle. For the most part, however, states have used case law to guide decisions about child residence and access.

Regardless of the availability of identified criteria or case law used to normalize the best interests of the

child principle, the lack of empirical research has limited its application. Namely, without empirical research as a guide, the application of the best interests of the child principle has been fraught with difficulty and fodder for individuals involved in residence and access disputes, formerly known as custody disputes, to manipulate its intent for their purposes. In particular, individuals involved in high-conflict divorces, often centered on children and assets, express the needs of their children predominantly in terms of their own needs.

High-Conflict Divorces

Although divorce presents challenges for all children, particular family dynamics associated with high-conflict divorces have been shown to influence child outcomes. How and to what degree of success parents negotiate their continued relationship has been found to have a significant impact on young children's post-divorce adjustment that is quantitatively and qualitatively different from other divorce factors. Within this renegotiation, the way parents communicate with each other, especially the conflict resolution strategies they use, is particularly important. Because of its reciprocal nature, these communication patterns include both communication behavior and interpretation of verbal and nonverbal cues by the other parent.

At the same time, divorce is often considered to be contentious by nature. Although some divorces are amicable, many are not. Thus, it is important to establish what constitutes a type of communication and a level of conflict associated with negative postdivorce child adjustment. Low levels of conflict and high levels of communication optimally meet the best interests of the child principle observed by the legal community for child residence and parental access decisions. The reverse has also been found to be true, where high levels of conflict in families with shared parenting responsibility result in poor adjustment outcomes for children. In particular, children of high-conflict divorce display behavioral, cognitive, emotional, and social problems to a greater degree than do other children from separated or divorced families.

Typical Child Effects

Typical child effects include feelings of sadness, anger, and fear, which may manifest in behavior changes and/or emotional and behavioral problems. It is common for these problems to become evident in the school setting. Typical in school-age and adolescent children is the desire or fantasy for parents to reunite, fear of abandonment by one or both parents, and/or feelings of responsibility for the divorce.

Infant/Toddler Age

Children express their difficulties in the area where they are currently developing. Several developmental theories suggest that infant and toddler-age children experience significant effects of divorce. Common behavioral manifestations in infant/toddler-age groups are increased crying; intensity of problems separating from either or both parents; regression to an earlier developmental stage, especially in toileting; and sleep problems.

Attachment is considered the most significant socio-emotional developmental milestone for the infant/toddler-age group. According to attachment theory, it is important to build an attachment to a primary caretaker. Thus, the risks associated with divorce for this age group are related to attachment formation. These risks include feelings of loss of contact with the primary attachment figure, resulting in child depression and regressive behavior, as well as separation and relationship difficulties in later development. The risks inherent to separation from the primary caretaker for extended periods may affect current development and also threaten the foundational task of attachment formation and subsequent developmental tasks. Attachment theory assumes that the quality of attachment is closely linked to caretaker sensitivity to the child's needs. One concept related to this assumption is the notion of the effect of these interactions over time in development of internal working models of attachment. Essentially, relationships with caretakers greatly influence the child's way of thinking about new people and experiences. Thus, attachment quality is considered an important link to later development.

School Age

The majority of research on divorce effects in children has been aimed at child adjustment in school-age children. These studies have commonly focused on cognitive effects, such as school performance, and psychological and social functioning, including behavioral problems. At this age, cognitive ability increases, including the ability to take another's

perspective. More specifically, the recognition of other people's feelings and goals, a more sophisticated type of reciprocity, is closely intertwined with cognitive ability in perspective taking. As a result, a couple of common behavior manifestations present themselves during this period for school-age children of divorcing parents. School-age children often feel anxious out of concern over loyalty to their divorcing parents. Also common at this age are feelings of anger that come with the lack of control they have over their current situation. The initiation of school-home communication and newly presented homework responsibilities pose a unique and additional challenge for these children as they frequently move between both parents' homes.

Adolescent Age

Adolescent children experience many of the same emotions that younger age groups do, but manifest them in different ways. One of the recognized adolescent milestones is the expression of a need for independence. Within the context of experiencing divorce within their families, adolescents may exhibit an effort to control their situation by demanding and then changing living arrangements with each parent. Because of the tendency toward risk-taking behavior in this age group, this can present challenges in potential negative risk-taking behaviors.

School Supports

Families embroiled in high-conflict divorces as well as typical divorcing families may be overwhelmed with both logistic and psychological challenges associated with divorce. In this way, school personnel educated about typical and age-specific divorce effects can support the detection of divorce effects and aid in intervention. Although all children experience divorce differently and in their own time frame, as has been noted, there are typical divorce effects found in children from divorcing families. These effects include feelings of sadness, anger, and fear, which may manifest in behavior changes and/or emotional and behavioral problems, and it is common for such behavior to become evident in the school setting.

School supports include adult and peer support and support of home-school communication. Adult support can be provided in the form of school counselors or other trained professionals. These individuals can help children to understand the divorce process and develop coping skills. As divorce becomes a typical social stressor for school-age children, schools are more frequently providing peer-based programs for children from divorcing/divorced families. Facilitated peer support groups have the effect of normalizing the divorce process for children. This can help children to reestablish a sense of personal control. Additionally, school personnel, including teachers, can support children from divorcing/divorced families in managing home-school communication. This type of support can be provided by way of considering binuclear families' special needs, such as when children go between two homes, as they pertain to home-school communication. This is a particularly important support to the child in the early stages of divorce, when levels of interparental conflict are typically higher and levels of positive interparental communication are typically lower.

Shannon Altenhofen and Zeynep Biringen

See also Emotional Development; Family Influences; Self-Concept; Self-Esteem

Further Readings

Amato, P. R. (2001). Children of divorce in the 1990s: An update of the Amato and Keith (1991) meta-analysis. *Journal of Family Psychology*, *15*, 355–370.

Amato, P. R., & Keith, B. (1991). Parental divorce and the well being of children: A meta-analysis. *Psychological Bulletin*, *10*, 26–46.

Lewis, J. M., Johnson-Reitz, L., & Wallerstein, J. S. (2004). Communication in divorced and single-parent families. In A. L. Vangelisti (Ed.), *Handbook of family communication* (pp. 197–214). Mahwah, NJ: Lawrence Erlbaum.

Sillars, A., Canary, D. J., & Tafoya, M. (2004). Communication, conflict, and the quality of family relationships. In A. L. Vangelisti (Ed.), *Handbook of family communication* (pp. 413–446). Mahwah, NJ: Lawrence Erlbaum.

Domestic Violence

The ways in which domestic violence is defined vary as much as the people who are the perpetrators and victims of domestic violence. Broadly speaking, domestic violence can refer to any verbally or

physically aggressive act committed within the context of a close interpersonal or familial relationship. This definition would include violence perpetrated from a husband against a wife, an adult child toward his or her mother, or even aggression between siblings. However, domestic violence more commonly refers to physical or psychological aggression committed by an intimate partner such as a husband, wife, girlfriend, or boyfriend against another intimate partner. Because of the prevalence of domestic violence in society, professionals are increasingly being confronted with the problem in schools, the legal system, and the workplace.

Prevalence

Estimates of the prevalence of domestic violence vary for a number of reasons. The violent acts that are frequently considered domestic violence can range from psychological abuse to sexual abuse. Psychological abuse typically refers to such things as insulting, swearing, or yelling at one's partner; stomping out of a room; saying something to spite one's partner; calling one's partner names; destroying something that belongs to one's partner; or threatening to hit a partner or throw something at him or her. Moderate levels of physical abuse can include actually throwing objects, twisting a partner's arm or hair, pushing, shoving, grabbing, or slapping. Severe physical abuse is normally characterized by use of a knife or gun, hitting with something that could hurt, choking, slamming against a wall, striking with a closed fist, kicking, or sexually assaulting a partner. The prevalence of domestic violence generally increases when less severe acts are included in the definition and decreases when only the most severe forms of domestic violence are included. Prevalence estimates also can vary because of the people who are sampled. Domestic violence occurs more frequently in couples who are unsatisfied in their relationship, couples who seek counseling, and among women seeking assistance at domestic violence shelters. Estimates also will vary depending on whether they are based on self-reports given over the phone, offered anonymously in a research setting, or based on actual crime statistics.

As a result, estimates of the prevalence of domestic violence vary across different groups. About one quarter of women report experiencing some form of physical or sexual violence at the hands of a romantic partner at least once in their lifetime. Probably about

10% of women will experience some form of severe violence, and women are more likely to be murdered by an intimate partner than by anyone else. Furthermore, more than 10% of women who are married or cohabiting report being physically abused annually, approximately 1 million each year. Of course, domestic violence is not only confined to married or cohabiting couples. The majority of high school and college students report being involved in an aggressive or violent relationship. Domestic violence also is not only confined to male-female relationships. Domestic violence occurs in heterosexual couples and homosexual couples, and evidence is increasingly suggesting that women also perpetrate domestic violence against men. Although domestic violence is not confined to heterosexual couples where the man is the perpetrator and the woman is the victim, most of the research focuses on these types of relationships. As a result, this discussion also primarily focuses on the male perpetration of domestic violence in heterosexual relationships unless expressly noted otherwise.

No matter the precise figures regarding the prevalence of domestic violence or even the nature of the relationship, it is clear that there are significant societal costs associated with domestic violence. A significant percentage of emergency room visits by women are the result of domestic violence, and women in abusive relationships are more likely to visit their family physician for a host of physical complaints, many of which are directly related to the abuse. Domestic violence calls are among the most common job-related tasks for law enforcement officers across the country. Men who abuse their partners also are more likely to abuse their children and in turn negatively affect their children's socialization and academic performance. Nonetheless, domestic violence cuts across all occupations, races, and socioeconomic groups. No person is immune from being a victim of domestic violence.

Risk Factors

Although domestic violence cuts across different groups, there are factors that increase the risk for the perpetration of domestic violence. These *risk factors* can be grouped into demographic, psychological, and contextual variables. The most relevant demographic factors typically include variables such as age, socioeconomic status, and race.

In general, the younger the perpetrator of domestic violence is, the lower the risk for committing

domestic violence. These age-related declines are even greater for more severe violence compared to moderate levels of domestic violence. Even though the wealthy, middle class, and poor are all victims and perpetrators of domestic violence, the lower one's socioeconomic status, the greater the risk for domestic violence. This increased risk is normally related to the additional financial and social pressures as well as the reduced social support and education more likely to occur in lower socioeconomic groups. The role of race and ethnicity has a complex relationship as a risk factor for domestic violence. Normally, perpetrators who are African American, Latino/Hispanic, or generally of minority status are at greater risk to perpetrate domestic violence. However, it is important to examine other factors that might be the cause of this relationship. Specifically, socioeconomic status appears to be related to race and ethnicity and may be the real reason for the increased risk within many minority groups. It is not simply that many minority groups are more likely to perpetrate domestic violence, but that they are more likely to be unemployed, live in poverty, be less educated, and work in lower-status occupations that are all related to a reduced socioeconomic status. When socioeconomic status is factored into the equation, race is not related to the perpetration of domestic violence.

Another category of risk factors commonly identified for the perpetration of domestic violence is psychological risk factors. Obviously, men who are more hostile and angry are more likely to perpetrate domestic violence. Surprisingly, though, men who are less assertive are also more likely to be violent. Assertiveness should not be confused with aggression. Assertiveness is the ability to appropriately express your feelings or ideas in a social situation, especially when those feelings or ideas may be met with resistance or are in the form of a request. Men who are more likely to perpetrate domestic violence have difficulty expressing themselves in an appropriately assertive manner and therefore often resort to violence or aggression out of the resentment they let build up. A lack of self-esteem is often related to poor assertiveness and in turn is also a risk factor for domestic violence perpetration. Another risk factor for domestic violence that will be examined more closely later is the need for power and control. Traditionally, treatment programs for domestic violence perpetrators focus on the need of the perpetrator to dominate and control his victim, and violence is one mechanism for him to do so. In addition to these

less pathological psychological risk factors, there also is a relationship between domestic violence and particular psychological disorders. Depression and other personality disorders, such as borderline personality disorder and antisocial personality disorder, have been consistently linked to domestic violence. An even stronger risk factor for domestic violence is the use and abuse of drugs and alcohol. Evidence suggests that not only are perpetrators of domestic violence more likely to suffer from a substance abuse disorder, but they are more likely to abuse alcohol or drugs immediately prior to the abuse itself.

The final category of risk factors related to domestic violence is the group of factors related to the relationship or the environmental context in general. Overall, it is very clear that perpetrators of domestic violence are more likely to be dissatisfied in their romantic relationship and that relationship dissatisfaction may be one of those most significant risk factors for domestic violence. This dissatisfaction may be evident in more disagreements or bickering or simply an emotional distance between the couple.

Another variable that may be somewhat surprisingly related to an increase in the severity of domestic violence perpetration is pregnancy. There may be a host of reasons for this risk factor, including additional dependency of the mother on the father, additional stress and financial commitments, and insecurity on the part of the father because of the change in their relationship. Other significant risk factors for the perpetration of domestic violence are related to childhood experiences. A child who is the victim of child abuse or witnesses domestic violence between his parents is at greater risk for the perpetration of domestic violence. This relationship is sometimes referred to as the intergenerational transmission of violence. The intergenerational transmission of violence refers to the fact that children who live in abusive relationships are more likely to perpetrate abuse as they become adults, enter into their own relationships, and thereby expose their children to abuse, which continues the cycle of violence.

Risk Assessment

Identifying risk factors for domestic violence is important for the assessment, prediction, and prevention of domestic violence. The term *risk assessment* refers to the formal evaluation of the potential risk that a perpetrator of domestic violence poses to

commit future domestic violence. Although risk assessments used to be satisfied with predicting the probability of future domestic violence, they are now concerned with much more than just the likelihood of future violence. Risk assessments also should evaluate other aspects of the future violence, such as the potential timing and severity of the violence. Risk assessments should rely on multiple sources to collect this information. It is not sufficient to rely only on the perpetrator for information but instead legal records, witnesses to the violence, and mental health records should be consulted. Most importantly, the risk assessment should be victim informed. The victim should be an integral component of any risk assessment of the perpetrator because she is likely to have a perspective that no other single person can provide. Finally, risk assessments should focus not only on assessing risk but also on managing it. *Risk management* refers to the process by which an attempt is made to reduce or avoid domestic violence. For example, part of the risk management process may be to refer the victim to a domestic violence shelter or encourage the perpetrator to seek psychological treatment for his domestic violence problem. Treatment for both the perpetrator and the victim are central to effective risk management.

Treatment and Intervention

A variety of interventions have been designed to manage or treat domestic violence. However, no single intervention will prevent all domestic violence or even most violent incidents. The best intervention is one that is multimodal or uses multiple approaches to solving the problem of domestic violence. Accordingly, the prevention and cessation of domestic violence can be accomplished only through the use of psychoeducation, community involvement, and the criminal justice system.

Psychoeducational Interventions

Psychoeducational interventions tend to focus on the men as perpetrators or women as victims, or on couples. The more traditional psychoeducational approaches to domestic violence focus on men as perpetrators. Typically, these programs are either feminist oriented or cognitive-behavioral, or a combination of the two. Traditionally, treatment programs have had a feminist orientation that encourages men to examine

their adoption of certain patriarchal messages within society. For example, the feminist model suggests that the cause of domestic violence is the power differential between men and women that modern society supports. The approach argues that men commit domestic violence in an attempt to dominate and control their female partners, as they have grown up in a society that encourages women to adopt a subservient role. The feminist model argues that only through recognition of this role and a rejection of it can domestic violence stop. The cognitive-behavioral model tends to focus on anger control and suggests that men perpetrate domestic violence largely because of certain aggressive supportive thought patterns they exhibit. For example, a perpetrator may think that his wife *never* listens to him, and that he can either continue to be ignored or become violent so she listens to him. Cognitive-behavioral approaches get perpetrators to identify these errors in their thinking and teach them to replace the thinking errors with more appropriate thoughts that then reduce the likelihood of being aggressive. It is not currently clear that either one of these approaches is superior to the other, but this finding may be because few batterer programs are strictly feminist based or strictly cognitive-behavioral and instead are a combination of the two.

Batterer Typologies and Treatment

As greater attention has focused on treatment interventions for men who perpetrate domestic violence, researchers have increasingly concluded that there is no single type of batterer. The batterer literature generally suggests that domestically violent men can be categorized into subgroups based on individual characteristics such as personality/pathology and the nature of their past violence. It has been suggested that these *typologies* could be useful in designing treatments and identifying the causes of domestic violence. For example, scholars have speculated that batterers with antisocial personality disorder and a history of generalized violence will be the least likely type of perpetrator to benefit from existing forms of treatment. Consistent with this speculation is that batterers who drop out of treatment are more likely to have drug and alcohol problems, and antisocial or narcissistic tendencies. They are also more likely to have committed more severe levels of domestic violence, and their probability of reoffending is elevated. As a result, it is believed that treatment programs for

men that focus on the specific needs of these different types of batterers are likely to improve treatment outcomes and increase the effectiveness of domestic violence treatment for men.

There are also psychoeducational programs designed to treat women, both as victims and as perpetrators. Traditional approaches focus on treating women as the victims by addressing issues related to self-esteem, removing any self-blame or guilt they might experience by incorrectly believing that they are the reason for the violence, and increasing their awareness about alternatives to the abusive relationship. Many people question why women remain in these violent relationships. The reasons are multiple and varied. Women in these relationships are frequently dependent on their partner for financial reasons and unable or unclear how to provide for themselves financially. In addition, the woman may have the added responsibility of being the primary caregiver for the children and may not know where to turn for child care assistance even if she could find employment. Perpetrators also attempt to isolate their victims away from their friends and family and therefore make escaping the relationship even more difficult. The psychological and physical abuse also takes a toll on the victims' self-esteem and belief that there is anything or even anyone outside of the perpetrator for them.

Battered Woman Syndrome

Lenore Walker originally used the term *battered woman syndrome* to explain the abusive relationship and the reasons why many women have difficulty leaving their batterer. Based on her observations with battered women, Walker believed that the violence occurred in cycles of three stages. During the *tension-building* stage, the abuse is largely verbal or very mild physical violence. The victim may attempt to appease the abuser and minimize the seriousness of the abuse. The abuse escalates and becomes the most severe during the *acute battering* stage. The acute stage is followed by *loving contrition.* During this stage, the abuser is profoundly apologetic for his behavior. He may completely change his attitude, shower the victim with attention and gifts, and promise never to be abusive toward her again. Walker believed that women who were repeatedly abused developed *learned helplessness*, based on the work of Martin Seligman. Seligman demonstrated that if dogs were shocked repeatedly without any way to escape the

shocks, they would eventually no longer try to escape even if a way for them to escape the shocks became obvious. Walker believed that the uncontrollable cycle of violence was similar to the inescapable shocks for the dogs. However, besides Walker's observations, there is very little support for the notion that all or even the majority of women in abusive relationships experience the cycle of violence or learned helplessness. Nonetheless, it remains a popular notion among domestic violence advocates, and battered woman syndrome has been used in court as a defense for women who kill their abusers.

Couple Treatment

Two more controversial treatment approaches for domestic violence are treating couples and treating women as the perpetrators of domestic violence. In couples' therapy, the partners focus on each of their contributions to the escalation of the violence when a discussion begins in a calm and rational manner. Although research has found this approach to be effective in some instances, there are many reasons it remains controversial. Some argue that by focusing on each partner's contributions to the violence, it removes responsibility from the real culprit, the perpetrator of the domestic violence, and places blame on the victim, further victimizing her. In addition, some believe that it is inappropriate to treat the perpetrator and the victim together because it encourages the violence and heightens the risk to the victim because she may say something in therapy that she is later punished for outside of the therapy session. Supporters of the approach reply that couples are not suitable for this intervention if there continues to be physical abuse in the relationship.

Interventions that focus on the woman as the perpetrator of domestic violence are also increasingly being used. These treatment approaches typically are cognitive-behavioral in orientation and tend to focus on anger control and other issues, such as trauma, that may be unique to women who perpetrate domestic violence. These approaches obviously conflict with the traditional feminist approach to treatment that suggests that only men can be the perpetrators because of the power differential that exists in our society. The notion that women perpetrate domestic violence is based on more recent but consistent findings in the literature that not only do some women perpetrate domestic violence, but also the perpetration rates between men and women may be equal and the

causes of their domestic violence similar. However, critics argue that even if there are equal rates of perpetration of domestic violence, the violence is not equal in any other aspect.

Many arguments exist that attempt to explain these equal rates of perpetration. First, critics argue that the equal perpetration rates are based largely on self-report data, and the research tends to suggest that men underreport their violence more than women. It also is argued that the violence typically perpetrated by women tends to be in self-defense. Women who act out in violence do so because they are being abused or are about to be abused, and they are attempting to retaliate while they are still physically capable. Another argument is that even if the frequency of the violence is equal, the severity of the violence is not. Men are more likely to perpetrate more severe violence that results in medical attention or greater physical injury because of their greater physical size and strength. Finally, there is a difference between the perpetration of domestic violence and *battering*. Critics argue that women may be able to perpetrate domestic violence, but they are not able to batter. Battering is usually distinguished from perpetration in terms of the severity of physical damage, the ability to socially and financially isolate the victim, and the general difference in the ability to dominate and control the victim. The argument is that men are able to batter women and that the negative effects of battering go beyond physical abuse.

Community-Level Interventions

In addition to the psychoeducational interventions for men, women, and couples, it is also important to have community involvement. Community involvement normally focuses on the use of advocates, shelters, and safe homes. Advocates are professionals who may work for organizations such as the YWCA or local domestic violence councils, or as a part of local government such as a prosecutor or police. Advocates typically assist the victims of domestic violence in navigating the system. This process may mean supporting the victim emotionally and encouraging her that she is doing the correct thing in seeking help for herself and her children, and in leaving her husband. It may mean talking to the victim about the difficulties that are likely to occur in filing criminal charges and prosecuting her partner for assault or domestic assault. Advocates also help victims identify the resources available to them in the community for providing basic necessities and protection if they have left their abusers and were financially dependent on them. One of those resources may include a safe house. Safe houses are places where women and their children can find temporary shelter. They often stay with individual families who are willing to take in women or consist of houses in secret locations so that the batterer cannot find the victim. Shelters tend to be more formalized programs that provide shelter and necessities for the victim and sometimes her children. Shelters may also provide psychological counseling and occupational counseling to assist the woman in living independently.

Criminal Justice Interventions

Another element that is important in preventing domestic violence is the criminal justice system. There are a number of ways in which the criminal justice system can take an active role in the prevention of domestic violence. First, law enforcement officers are increasingly being trained in the de-escalation of domestic violence and some of the subtleties involved in domestically violent relationships. Part of the investigative aspect of domestic violence as a crime is the use of *mandatory arrest polices*. Mandatory arrest policies are common in many major cities and counties. Mandatory arrest policies require an arrest if it appears an act of domestic violence has occurred. These policies were met with initial enthusiasm. It was believed that they would allow for a separation of the perpetrator and victim, result in a cooling-off period, give the victim a chance to escape, and encourage the prosecution of the perpetrator. However, research has failed to universally support the use of mandatory arrest policies, and some research suggests that domestic violence increases once the perpetrator returns. Other policies that can be implemented by prosecutors and judges are the use of restraining orders and no-drop policies. Restraining orders require perpetrators to refrain from any contact with the victim and to remain an identified distance away from the victim. No-drop policies mandate that the prosecution cannot drop the charges against the perpetrator and can even force the victim to participate, even when it frequently occurs that she is unwilling to participate after she has reconciled with the perpetrator.

Matthew T. Huss

See also Aggression; Child Abuse; Learned Helplessness

Further Readings

Archer, J. (2002). Sex differences in physically aggressive acts between heterosexual partners: A meta-analytic review. *Aggression and Violent Behavior*, 7, 313–351.

Babcock, J. C., Green, C. E., & Robie, C. (2004). Does batterers' treatment work? A meta-analytic review of domestic violence treatment. *Clinical Psychology Review*, 23, 1023–1053.

Dutton, D. G. (1998). *The abusive personality: Violence and control in intimate relationships.* New York: Guilford.

Holtzworth-Munroe, A., Meehan, J. C., Herron, K., Rehman, U., & Stuart, G. L. (2000). Testing the Holtzworth-Munroe and Stuart typology. *Journal of Clinical and Consulting Psychology*, 68, 1000–1019.

Holtzworth-Munroe, A., & Stuart, G. L. (1994). Typologies of male batterers: Three subtypes and the differences among them. *Psychological Bulletin*, 116, 476–497.

Huss, M. T., & Langhinrichsen-Rohling, J. (2000). The identification of the psychopathic batterer: Clinical, legal, and policy implications. *Aggression and Violent Behavior*, 5, 403–422.

Langhinrichsen-Rohling, J. (2005). Top 10 greatest "hits": Important findings and future directions for intimate partner violence research. *Journal of Interpersonal Violence*, 20, 108–118.

Langhinrichsen-Rohling, J., Huss, M. T., & Ramsey, S. (2000). The clinical utility of batterer typologies. *Journal of Family Violence*, 15, 37–53.

Langhinrichsen-Rohling, J., Huss, M. T., & Rohling, M. (2006). Assessing violence in adults. In M. Hersen (Ed.), *Clinician's handbook of adult behavioral assessment* (pp. 371–400). New York: Elsevier.

Sartin, M., Hansen, D., & Huss, M. T. (2006). Recidivism and treatment response-related characteristics for domestic violence perpetrators. *Aggression and Violent Behavior*, 11, 425–440.

Straus, M. A., Hamby, S. L., Boney-McCoy, S., & Sugarman, D. B. (1996). The revised Conflict Tactics Scales (CTS2): Development and preliminary psychometric data. *Journal of Family Issues*, 17, 283–231.

DRUG ABUSE

The definition of drug abuse has fluctuated with changes in cultural and legal norms, and there is currently no consensus on a standardized definition. However, drug abuse generally can be defined as the harmful or risky use of legal or illegal drugs or the use of legal drugs in a manner or amount inconsistent with medical advice. Some organizations such as the American Psychiatric Association (APA) do not refer to the abuse of drugs per se, but rather to the abuse of substances in general. These two terms (*drug abuse* and *substance abuse*) are often considered synonymous and are used synonymously in this entry.

The abuse of drugs is frequently associated with a myriad of psychosocial issues, including violence, homicide, homelessness, premature death, psychopathology, and economic loss. For secondary school-aged and college-aged individuals in particular, drug abuse is also correlated with school failure, criminal activity, accidents, aggressiveness, unsafe sex, unplanned pregnancy, and suicide.

Humankind has used alcohol and other drugs in one form or another since the beginning of recorded history. Alcohol was consumed as early as the Paleolithic era, and other drugs such as cocaine, marijuana, and opium have all been used in the past for religious or medical purposes. Even today, the abuse of both legal and illegal drugs is relatively commonplace despite moral implications and legal repercussions. Drug abuse is both a major public health issue and a highly pertinent matter facing educators today. Given the reliable association between drug abuse and impaired academic performance (among other variables relevant to students), educators, researchers, and practitioners of educational psychology need to be acutely aware of the following in order to help maximize the performance and adjustment of individuals in educational settings: the costs of drug abuse, trends in drug abuse, gender differences, theories of drug abuse, risk factors for drug abuse, the diagnosis and assessment of drug abuse, treatment options, and issues in prevention. These topics are addressed in this entry.

Costs and Trends of Substance Abuse

Economic Costs

The economic costs associated with alcohol and other drug abuse are staggering. According to the Office of National Drug Control Policy (ONDCP), the overall cost of drug abuse other than alcohol in 2002 in the United States was in excess of $180 billion. The National Institute on Alcohol Abuse and Alcoholism (NIAAA) also estimated that, in 1998, the cost related to alcohol abuse alone was also in excess of $180 billion. In both of these cases, the majority of costs were attributed to lost productivity and health care costs. The ONDCP also reported that in the year 2000, Americans spent an estimated $36 billion on

cocaine, $11 billion on marijuana, $10 billion on heroin, $5.4 billion on methamphetamine, and $2.4 billion on other illegal drugs.

Trends in Use and Abuse

The Substance Abuse and Mental Health Services Administration (SAMHSA) conducts an annual National Survey on Drug Use and Health (NSDUH) of Americans aged 12 years or older, and this is the primary source of information on the use of illicit drugs, alcohol, and tobacco in the civilian, noninstitutionalized population of the United States. According to the results of the 2005 NSDUH, an estimated 19.7 million Americans (or 8.1% of the population) reported using an illicit drug during the month prior to the survey. The most commonly used illicit drug was marijuana (14.6 million users in the past month), followed by prescription drugs used nonmedically (6.4 million), cocaine (2.4 million), and hallucinogens (1.1 million). Slightly more than 50% of Americans aged 12 years or older had consumed alcohol within the 30 days prior to the survey, which translates into about 126 million Americans.

The prevalence of drug abuse in youth should be of special interest to the educational psychology professional, considering that the current rate of illicit drug use among youths aged 12 to 17 within the past 30 days is approximately 11%, that approximately 28% of individuals between the ages of 12 and 20 drank within the past month, and that nearly 19% of them engaged in a binge-drinking pattern (defined as at least five drinks in a day). Researchers at the University of Michigan conduct an annual study on drug use trends among 8th-, 10th-, and 12th-grade students called the Monitoring the Future (MTF) study. Each year, the MTF study gathers self-report survey data from approximately 50,000 students to determine drug use rates and attitudes. According to the 2005 MTF study, alcohol is cited as the most commonly abused drug, followed by cigarettes and marijuana. This trend in alcohol abuse is especially troubling, considering that recent National Institutes of Health (NIH) studies have revealed that young people who began using alcohol before age 15 are four times more likely to develop alcohol dependence during their lifetime than those who began drinking at age 21 or later. Moreover, these youth are seven times more likely to report having been in an alcohol-related traffic accident during adolescence and adulthood.

Nevertheless, in 2006, the MTF study found a 23% reduction in the number of adolescents using illegal drugs between 2001 and 2006, including decreases in the use of marijuana, methamphetamine, steroids, LSD, and ecstasy. The study also found that alcohol and tobacco use had declined 13% and 29%, respectively. Conversely, the study also found that prescription drug abuse among young people had increased substantially during the same 5-year period. For example, past-year use of oxycodone, a prescription opioid, for nonmedical purposes had almost doubled since 2002. This trend has also been observed by several recent studies in both Canada and the United States, which report that the use of prescription opioids has become the predominant form of illicit opioid use. Although alcohol, tobacco, and marijuana remain the most commonly abused drugs, prescription drug abuse is quickly rising.

Gender Differences

Based on the results of the 2005 NSDUH, men reported being current drinkers more frequently than women. However, among the 12–17 age group, rates of current alcohol consumption were actually slightly higher for female youth in comparison to male youth (17.2% vs. 15.9%, respectively). This finding is consistent with the 2005 MTF study, which found that although males generally have relatively higher rates of heavy drinking than females, this pattern was reversed for 8th-grade and 10th-grade students.

In speaking specifically to alcohol use, women are more likely to be solitary drinkers and to drink in the home. It has also been consistently found that, on average, women tend to start drinking alcohol later in life, to drink alcohol less than men, and to have less adverse consequences related to alcohol use. However, in general, the same amount of alcohol or another drug taken by a man and a woman will have a stronger effect on the woman. The main reason for this is the biological tendency for women to have a higher percentage of body fat and a lower percentage of body water. Because women have proportionally more fat than water, there is less dilution of the drug while it is in the body of a woman. In addition, certain drugs bind selectively to fat cells in the body, which further contributes to their slower metabolism and elimination from the body of a woman.

Based on the 2005 NSDUH (as well as the 2005 MTF survey), males age 12 and older were more

likely to report current illegal drug use in comparison to females, particularly with respect to marijuana, whereas rates of using prescription-type drugs for nonmedical purposes were almost equal across gender. Among youth between the ages of 12 and 17, rates of illegal drug use were found to be very similar across gender, whereas the other two general trends still prevailed (i.e., male youth report more marijuana use, and abuse of prescription-type medications was approximately equal).

According to the National Institute for Drug Abuse (NIDA), although men are more likely than women to find themselves in situations where drugs are available, given the opportunity to use drugs for the first time, both are equally likely to do so. Men and women are also equally likely to have problems with cocaine, heroin, hallucinogens, tobacco, and inhalants. However, there appears to be gender differences in vulnerability to abusing specific substances. Women are more likely than men to have problems with sedatives, hypnotic drugs, and drugs designed to alleviate anxiety, whereas men are more likely to abuse marijuana and alcohol.

Theories of Substance Abuse

Although there is still widespread dispute about the cause of drug abuse, several models have gained prominence. The moral, disease, and biological models argue that the cause of drug abuse originates from within the individual, whereas social learning theory, systems theory, and the sociocultural model view the social environment as most influential. The biopsychosocial model combines elements from several of these theories into a single, integrated approach.

Individual Theories

The moral model argues that drug abuse is a choice, and that individuals are personally and morally responsible for the problems that may result from drug abuse. Alternatively, the disease model postulates that drug abuse is a progressive and irreversible disease with biological origins and perhaps also psychological and spiritual causes. This model is the foundation of many self-help groups such as Alcoholics Anonymous and Narcotics Anonymous. The biological model, which appears to be endorsed by the National Institute on Drug Abuse (NIDA), posits that drug abuse is largely a result of genetic and physiological processes. NIDA argues that although initial drug use may be voluntary, the ensuing abuse is often driven by neurological changes that affect behavior.

Social Theories

Social learning theory, first proposed by Albert Bandura, when applied to drug abuse presumes that drug abuse results from complex learning, including modeling and other forms of observational learning, and is grounded in the interaction of the individual with his or her environment. In contrast, systems theory places greater primacy on familial processes, such as dysfunctional interaction patterns within a family system, that can lead to drug abuse. The sociocultural model proposes that drug abuse is a logical consequence of the particular social norms and laws of many subcultures and societies.

Integrated Approaches

The biopsychosocial model integrates three major domains known to influence behavior: biological, psychological, and social factors. The biopsychosocial model posits that each of these factors contributes in some interactional way to drug abuse. Although this model is more complex than the other major theories, many argue that it more accurately reflects the complexity of human experience.

Another integrated approach developed by James Prochaska, John Norcross, and Carlo DiClemente is the transtheoretical model of change. It is an evidence-based model used to guide assessment, treatment, and policy, and is discussed in further detail in the diagnosis and assessment section.

Risk Factors for Drug Abuse

Although research has identified a number of risk factors for drug abuse, there is no single, reliable predictor of who will abuse drugs and who will not. With this caveat in mind, the Center for Substance Abuse and Prevention, a branch of the Department of Health and Human Services' SAMHSA, has organized risk factors for drug abuse into three groups: individual factors, family factors, and environmental factors. Individual risk factors include antisocial behavior (such as criminal activity and violence), social isolation, high need for independence, psychopathology, positive attitudes toward drug use, poor impulse control, and thrill-seeking

personality type. Family factors include parental drug use, poor family relationships, interpersonal conflict, and physical or sexual abuse. Environmental factors include peer influence, rejection or low acceptance, cultural or social norms, poverty, neighborhood disorganization, scholastic failure, low scholastic involvement, and negative school environment. Awareness of these risk factors can help in the early identification of individuals likely to begin abusing or continue abusing drugs.

Diagnosis and Assessment

Diagnosing drug abuse is essential for the provision of treatment services, yet the abuse of drugs is often difficult to assess. Many individuals who abuse drugs are not willing to disclose their habit, may be plagued by feelings of shame, or may fear legal retribution for their disclosure. Moreover, individuals who abuse drugs may deny the seriousness of their problem, both to themselves and to others. In light of these barriers to effective assessment, in the book *Where to Start and What to Ask: An Assessment Handbook*, Susan Lukas has suggested that questions such as the following may be useful in helping determine whether an individual needs a formal assessment for drug abuse: What substances does the individual use? How recently has he or she used them? How much does he or she use? When does he or she use? What happens when he or she uses? Why does he or she use? What effect is the substance having on his or her life? Does he or she feel guilty about his or her use? Has the individual ever tried to cut down or stop using? How did he or she try? Did it work?

After these types of questions have been answered, seeking a diagnosis may be useful in determining what type of treatment will be most beneficial. The most widely used and accepted method of diagnostic classification in Canada and the United States is the *Diagnostic and Statistical Manual of Mental Disorders*, (Fourth Edition, Text Revision) (*DSM-IV-TR*), published by the American Psychiatric Association. The use of *DSM-IV-TR* diagnostic terminology is useful because it offers standardized, descriptive criteria that can be used to ensure effective clinical communication about drug abuse.

According to the *DSM-IV-TR*, substance abuse is essentially characterized by clinically relevant impairment or distress as evidenced by at least one of the following symptoms for a 12-month time frame:

(a) repeated failure to fulfill major work, school, or home obligations; (b) frequent use in physically hazardous circumstances (including driving under the influence of substances); (c) persistent substance-related legal troubles; or (d) recurrent use despite ongoing or repetitive social or interpersonal problems caused by or worsened by substance use. Categories of drugs that qualify for a diagnosis of substance abuse, according to the *DSM-IV-TR*, are alcohol; amphetamines; cannabis; cocaine; hallucinogens; inhalants; opioids; phencyclidine; and the category of drugs comprising sedatives, hypnotics, or anxiolytics (anxiety-reducing drugs). Drugs that do not fit these categories can be classified as "other" (i.e., abuse of other substances). The *DSM-IV-TR* further distinguishes between substance abuse and substance dependence, the latter of which is characterized by at least 12 months of the continued presence of three or more of the following symptoms: (a) drug tolerance; (b) drug withdrawal; (c) loss of control over use (in terms of amount or duration); (d) sustained desire or repeated unsuccessful attempts to stop or cut down; (e) substance-related preoccupation (e.g., most of one's time is spent obtaining the substance, using the substance, or recovering from its effects); (f) reducing or eliminating important social, occupational, or recreational activities; and (g) continued use despite evidence that a physical or psychological problem is caused or aggravated by substance use. The diagnosis of substance dependence is further specified by the presence or lack of physiological dependence (i.e., evidence of tolerance or withdrawal symptoms).

The International Statistical Classification of Diseases and Related Health Problems, Version 10 (ICD-10) is a diagnostic system developed by the World Health Organization that is used in many countries outside of Canada and the United States. The ICD-10 defines harmful substance use (similar to the *DSM-IV-TR* diagnosis of substance abuse) as a pattern of psychoactive substance use that results in physical or psychological damage to one's health. Like the diagnosis of substance dependence in the *DSM-IV-TR*, the ICD-10 includes criteria for diagnosis of dependence syndrome to describe drug addiction. Dependence syndrome is defined as a cluster of behavioral, cognitive, and physiological phenomena that develop after repeated substance use and that typically include (a) a strong desire to take the drug, (b) difficulties in controlling usage, (c) persistent use despite harmful consequences, (d) a prioritizing of drug use over other

activities and commitments, (e) increased drug tolerance, and (f) possible physical withdrawal symptoms.

Many agencies in the United States and elsewhere have implemented Prochaska, Norcross, and DiClemente's transtheoretical model as a method of assessment and treatment. It is an integrated, evidence-based model that describes a stage-based pattern of change based on an individual's readiness to change. According to the transtheoretical model, determining what stage of change an individual is in is essential to providing appropriate and effective treatment. The individual who abuses drugs is assessed to determine what stage of change he or she is in, and treatment strategies geared to that particular stage of change are then implemented. The stages are as follows:

- *Precontemplation:* The individual does not recognize the need for change and does not have any intention of changing his or her pattern of drug abuse.
- *Contemplation:* The individual is seriously considering making a change. However, the individual does not have specific plans to stop abusing drugs in the near future.
- *Preparation:* The individual is planning to stop abusing drugs within the next month, but may still be ambivalent about taking action.
- *Action:* The individual has stopped abusing drugs but has done so for less than 6 months.
- *Maintenance:* The individual has stopped abusing drugs for more than 6 months.
- *Relapse:* The individual has returned to abusing drugs.

Treatment

Treatment Types and Settings

Although some individuals who abuse drugs resolve their problem spontaneously without help, formal treatment has been shown to be helpful for many individuals. There is evidence that those who receive some type of professional help subsequently reduce their use of substances and show improvement in other areas of functioning. In addition, psychological treatments targeting significant life problems also tend to improve substance abuse recovery outcomes. Reliable evidence has been demonstrated for the effectiveness of a number of treatment approaches, including social skills training, self-control training, brief motivational counseling, behavioral marital therapy, stress management training, and community reinforcement

approaches. Some evidence of positive effect has also been found for covert sensitization, behavioral contracting, the medication disulfiram (brand name: Antabuse), and antidepressants.

Treatment can be implemented in a variety of settings, which may include inpatient medical detoxification and stabilization in acute care facilities, dual-diagnosis hospital inpatient care, short- and long-term residential treatment, outpatient full-day treatment, temporary transitional residential treatment (halfway housing), and intensive or nonintensive outpatient programs.

Best Practices Guidelines

Both Health Canada and the NIDA have independently developed a series of research-driven best practices guidelines for drug abuse treatment and rehabilitation. Some highlights of these best practices include the following:

- Pharmacotherapies such as methadone and disulfiram can be effective as supplemental treatments when used in a controlled setting.
- Behavioral therapies such as relapse prevention programs, behavioral self-control therapy, and behavioral contracting have been shown to be useful when administered either individually or in group settings.
- In terms of alcohol abuse, a community reinforcement approach has been shown to be effective for those with few social resources and relatively severe use.
- Marital therapy and social skills training are both well supported by research.
- Stress management interventions have been shown to be useful as part of alcohol treatment programs.
- Services should be flexible and individualized, and guidelines for the selection of appropriate services are essential. No single treatment is appropriate for all individuals, and treatment needs will likely change as treatment progresses.
- Group therapy is often preferable to individual treatment, unless otherwise contraindicated.
- Although research generally supports the relative cost-effectiveness of outpatient treatment, some individuals may benefit from both residential and outpatient or day programs.
- Effective treatment combines interventions that address the multiple needs of the individual, not just his or her drug use.
- Brief interventions of up to eight sessions have been shown to be especially effective for socially stable individuals with low to moderate alcohol dependence,

and are usually as effective as treatments of a longer duration for this group of individuals.

- Counseling provided by competent therapists with strong interpersonal skills is related to positive outcomes.
- The majority of those who have drug abuse problems do not seek help; therefore, more effort is required to increase awareness of specialized services to the general public as well as to health service providers.

It is expected that widespread application of these guidelines in drug abuse treatment will contribute to increased positive outcomes for substance-abusing individuals seeking support.

Prevention

Many people believe efforts at preventing drug abuse are vastly underfunded in comparison to the funds put toward treatment. Prevention in the context of drug abuse refers to the avoidance or mitigation of drug abuse and problems associated with drug use. Drug abuse prevention can involve primary, secondary, or tertiary prevention. The goal of primary prevention is to prevent drug use or abuse from occurring in the first place. Early education campaigns that encourage children to "just say no" would be an example of a primary prevention strategy. Secondary prevention involves interventions applied to those who are already in the early stages of using or abusing drugs in order to prevent the development of additional problems. An example of secondary prevention would be mandated alcohol education classes for those convicted of driving under the influence. Tertiary prevention involves reducing or stopping further deterioration among those with an established history of drug abuse. Harm reduction strategies designed to reduce harmful effects for those who continue to abuse drugs are seen as tertiary prevention strategies. An example of a harm reduction strategy is safe injection sites, such as those in Vancouver, British Columbia, Canada, that permit individuals to inject self-obtained illegal drugs in a setting staffed by medical professionals in an effort to prevent overdoses and the spread of communicable diseases.

Three major models of drug abuse prevention have emerged over the years, each with its own underlying philosophy and subsequent suggestions for prevention policies. The sociocultural model argues that social norms are the key factor in promoting the abuse of drugs. An example of the sociocultural model of prevention is social-norm education (as provided by many college health centers and college counseling centers), which seeks to counter common misperceptions about typical levels and patterns of drug usage. The consumption model argues that the prevalence of drug abuse is a direct function of the average levels of consumption in a given culture (e.g., the more the individuals in a culture drink, the more individuals in that culture who will abuse alcohol) and seeks to minimize the negative consequences of drug use through the implementation of restrictions. Higher taxes and age minimums for legal alcohol consumption would fall within the consumption model of prevention. The proscriptive model goes one step further and argues that if the availability of drugs (particularly alcohol) is completely prohibited, those who continue to abuse drugs can be regarded as bad or immoral. Therefore, the proscriptive model of prevention advocates for prohibition of availability and complete abstinence. The American Alcohol Prohibition Era (1921 to 1932) is an example of the proscriptive model at work.

In the past, several mass media campaigns have been aimed at youth and young adults in an effort to reduce drug abuse. Although these campaigns have been shown to increase public awareness and knowledge about drugs, their effectiveness at actually reducing drug use and abuse has not been consistently demonstrated. A recent prevention strategy that has shown promise is resistance skills training, which helps youth develop their problem-solving and decision-making skills. It also seeks to help youth develop cognitive skills for resisting media-based drug use messages, increase self-awareness and self-esteem, learn nondrug coping strategies, and develop interpersonal communication and assertiveness skills. The most widely used program of this kind in the United States is the Drug Abuse Resistance Education (DARE) program. Unfortunately, research on its effectiveness has yielded mixed results.

Robinder P. Bedi and Carlton T. Duff

See also *Diagnostic and Statistical Manual of Mental Disorders*

Further Readings

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., Text rev.). Washington, DC: Author.

Johnston, L. D., O'Malley, P. M., Bachman, J. G., & Schulenberg, J. E. (2006). *Monitoring the future national results on adolescent drug use: Overview of key findings, 2005* (NIH Pub. No. 06–5882). Bethesda, MD: National Institute on Drug Abuse.

Johnston, L. D., O'Malley, P. M., Bachman, J. G. & Schulenberg, J. E. (2006, December 21). *Teen drug use continues down in 2006, particularly among older teens; but use of prescription-type drugs remains high.* Retrieved January 14, 2007, from http://www.monitoringthefuture.org

Lukas, S. (1993). *Where to start and what to ask: An assessment handbook.* New York: Norton.

McLaughlin, T. F., & Vacha, E. F. (1993). Substance abuse prevention in the schools: Roles for the school counselor. *Elementary School Guidance & Counseling, 28,* 124–132.

Migneault, J. P., Adams, T. B., & Read, J. P. (2005). Application of the transtheoretical model to substance abuse: Historical development and future directions. *Drug and Alcohol Review, 24,* 437–448.

Rubin, E. (2003). Integration of theory, research, and practice: A clinician's perspective. In F. Rotgers, J. Morgenstern, & S. T. Walters (Eds.), *Treating substance abuse: Theory and technique* (2nd ed., pp. 343–363). New York: Guilford.

Schydlower, M., Anglin, T. M., Fuller, P. G., Jr., Heyman, R. B., Jacobs, E. A., et al. (1995). The role of schools in combatting substance abuse. *Pediatrics, 95,* 784–785.

Terry-McElrath, Y. M., Johnston, L. D., O'Malley, P. M., & Yamaguchi, R. (2005). Substance abuse counseling services in secondary schools: A national study of schools and students, 1999–2003. *Journal of School Health, 75,* 334–341.

Thather, D. L., & Clark, D. B. (2006). Adolescent alcohol abuse and dependence: Development, diagnosis, treatment and outcomes. *Current Psychiatry Reviews, 2,* 159–177.

Tolan, P., Szapocznik, J., & Sambrano, S. (Eds.). (2007). *Preventing youth substance abuse: Science-based programs for children and adolescents.* Washington, DC: American Psychological Association.

Vega, R. R., & Seligman, L. (2005). Diverse drug abusing populations. In R. H. Coombs (Ed.), *Addiction counseling review* (pp. 129–148). Mahwah, NJ: Lawrence Erlbaum.

Web Sites

Center for Substance Abuse Prevention: http://prevention.samhsa.gov/

Higher Education Center for Alcohol and Other Drug Abuse and Violence Prevention: http://www.higheredcenter.org

Office of Safe and Drug-Free Schools: http://www.ed.gov/about/offices/list/osdfs/

DYNAMICAL SYSTEMS

Educational psychology focuses on theories of learning that ultimately affect how students are taught. Behavioral learning theories influenced the teaching/ learning process for more than 50 years. In the 1960s, the information-processing approach brought the mind back into the learning process. The current emphasis on constructivism integrates the views of Jean Piaget, Lev Vygotsky, and cognitive psychology. Additionally, recent scientific advances have allowed researchers to shift attention to biological processes in cognition. The problem is that these theories do not provide an integrated approach to understanding principles responsible for differences among students in cognitive development and learning ability. Dynamic systems theory offers a unifying theoretical framework to explain the wider context in which learning takes place and the processes involved in individual learning.

Dynamic (or dynamical) systems theory is part of a paradigm shift involving the acceptance of chaos and complexity as theoretical frameworks in physics, biology, chemistry, engineering, ecology, and psychology. Its concepts can be applied to any systems that change over time. It emphasizes the complexity of learning and the students themselves. It recognizes the importance of context in change. It connects the physical with the mental. It allows teaching and learning to be connected processes.

Dynamical Systems Are Self-Organizing

Dynamic implies synergistic, changing, and chaotic (i.e., underlying order that appears random). *System* denotes an assemblage of interacting components whose essential properties arise from the relationships between its parts. Students, teachers, classrooms, schools, and school districts make up dynamical systems in education.

Students and teachers are individuals made up of diverse systems—biological, affective, and cognitive. There is also diversity based on race, ethnicity, gender, disability, and socioeconomic status—other systems that are an integral part of students and teachers. All of these systems contribute to individual differences in rate of learning, level of thinking, memory, and motivation. Pedagogically, teachers need to be aware of the dynamic complexity of their students and themselves

in order to more effectively understand what occurs in the teaching-learning process. For example, an African American student who has lived in a rural town will perceive information differently from an African American student who has been raised in an inner-city ghetto. Even though both students are African American, the differences in the systems that make up who they are will contribute to diverse outcomes. If both are learning the word *field*, their understanding of the meaning of this word will reflect their different experiences. The same teaching strategy may not work for both of them. For one, a trip to a field might be necessary whereas the other will have experience with fields and can easily make a connection. A view of students as dynamic systems emphasizes the individuality and complexity of all students.

Each dynamic system, in addition to being fluid and integrative, exists in a *state space*, which is an abstract construct depicting the range of behavior open to a system and constrained by the degrees of freedom available to the elemental components of the system. Each student has his or her own state space or possible behaviors. For example, a student whose family speaks Spanish at home is going to be constrained by English being a second language. His or her state space will be different from that of a student for whom English is a primary language. In order to be effective, a teacher needs to be sensitive to the different state spaces of his or her students. Some students may be incapable of carrying out certain behaviors because of the limits of their state space. The student with Spanish as the primary language may be incapable of understanding certain English idioms until he or she has learned other words first and changed his or her state space through self-organization.

Where a system is located in state space is often determined through *self-organization*, which occurs when a system is attracted to a preferred state of being out of many potential states. The individual components of a system emerge into new patterns (states of being) without anyone or anything directing this change. Self-organization emerges from the confluence of components within the system. In order for self-organization to occur, the system must be complex and open to changes in the environment.

Self-organizing, nonlinear dynamic systems tend to migrate toward certain conditions or behaviors that are called *attractors*. Attractors have varying degrees of stability. Some are easily changed, such as the heart rate when exercising. Others are not easily changed,

such as the stomach producing acid in response to food. An attractor state for a student might be his or her reading level, which may be stable for a long period of time.

Systems will remain in a certain state space until a perturbation pushes the system or increases the system's attraction to another state. *Perturbation* is a disturbance to the system. When disturbed, a system may move away from its present attractor and toward a new attractor, resulting in a *phase shift*. A phase shift is a new form that emerges from the loss of stability of the existing forms. For example, the student whose reading level has been stable for a time period might change to a higher level because of an interest in a more difficult book. The resulting phase shift might be observed as the student begins reading chapter books instead of picture books. There has been a phase shift in reading levels.

From a dynamic systems approach, learning is a self-organizational process on the part of the individual. The student undergoes phase shifts (transitions) in which the cognitive system self-organizes and new patterns of understanding emerge. Students pass from one organized state of the system to another. Along with the cognitive system undergoing change, neuronal networks are strengthened. As these systems change, they affect the other systems that make up the individual student.

Phase shifts can also be facilitated by a *control parameter*. A control parameter is a variable outside the system to which the system is sensitive and that moves the system through different states. Teachers use control parameters to perturb the attractor states of the students. In pedagogy, a control parameter is the teaching method. Some teaching methods are more successful than others. If successful, self-organization (learning) occurs. If there is no immediate self-organization, then the control parameter perturbation may disturb the system so that at some later point in time, self-organization may occur.

For example, if a teacher is focusing on addition in an elementary classroom, he or she will use methods to bring about an understanding of addition. When a student understands what addition is, he or she undergoes a transition. Not only has his or her cognitive system changed, but biologically, there are changes in the brain; emotionally, the student might feel satisfaction and pride about mastering something in school; and environmentally, the student might be able to use addition in new aspects of his or her life, such as shopping with parents. The student, as

a dynamic system, has self-organized and is no longer the individual he or she was prior to the new learning.

One truth about learning is that it evolves over time. Therefore, learning does not occur in every student each time a topic is covered in class. What often occurs is a perturbation of the attractor state. There can be many small perturbations without a change in the cognitive system. It is when the system (student) self-organizes that learning takes place. For example, a teacher may introduce addition to students in a number of different ways—direct instruction, small group activities, individual worksheets. The actual comprehension of what addition means will occur in individual ways. Some students may understand addition from the first activity, and therefore will self-organize, move to a new attractor state in which being able to add is the norm, and be able to demonstrate an understanding of addition. Other students may require all three strategies before meaning is created. Each activity perturbs the attractor state of these students, but no self-organization occurs until all activities have been experienced.

Dynamical Systems Are Nonlinear

Besides self-organizing, dynamic systems are also *nonlinear*, that is, output is not proportional to input. Teachers learn early that sometimes the best lesson plans do not bring about the learning that was sought. Pedagogically, there is not always a direct relationship between teaching and learning. Sometimes, a teacher does not have direct control over the specific learning that goes on, such as in a cooperative learning group, where students can learn from other students without the teacher's involvement. Most of the time, the teacher has to use a number of strategies to perturb the systems (students) until finally self-organization (learning) occurs.

All Systems Have a History

Dynamic systems theory espouses that all systems have a history that affects their states in the present. As the student acts, that action becomes part of his or her dynamical history. This history will then influence future actions. Fritjof Capra states that living structure is always a record of previous development. The history of a system includes its tendencies and constraints. Self-organization taking place in the present is constrained by self-organization of the past. The system's history will affect its present state space, attractors, phase shifts, and control parameters.

So, besides viewing students as dynamical systems, teachers must be aware that each of these systems (students) has a history that will affect his or her ability to learn. Some students have a history of doing well in school, understanding academic subjects, and having friends and supportive parents. Unless severely perturbed, these students will continue to do well in school throughout their school lives and respond well to teaching strategies. Other students do not have such a positive experience in their pasts. They may have problems being in school, may have difficulties learning academic subjects, and do not have supportive parents. These students may be harder to reach and require larger perturbations to their systems before learning can occur. Their histories may interfere with success in school. These are often the students who present the challenge to teachers.

Besides the students and teacher being systems with a history, the classroom community, too, is a system that has an ongoing, ever-changing history—a history of interactions within the enclosed walls of the classroom itself. No two days are alike in the classroom; indeed, no two moments in time are alike. The setup of the classroom—temperature, luminosity, desk orientation, ambient noise, odors—defines its state space. Within that broader framework, each member of the classroom community, each student and teacher, is an independent, complex, nonlinear system with a unique state space of his or her own and his or her own unique and constantly changing history. The variability and unpredictability of the classroom environmental system keep it in constant flux.

The environment also constrains the degrees of freedom available to individual students and teachers. For example, students learn to raise their hands to speak, speak softly or not at all while working, and get into line in the proper order. All of these constrain the behaviors of the students from what they might be if students did not have certain norms. Some students learn from this history of classroom behavior and automatically apply it in subsequent grades so that by the time the students reach later elementary years, the attractor state for classroom behavior is stabilized.

Understanding Dynamical Systems Within the Context of the Whole

Because a number of systems make up an individual, all of them, together, provide the essence of the

individual. Unlike the behaviorist reductionist view, which requires studying parts of the whole and then generalizing across the whole, dynamic systems theory purports that the whole is always different from the mere sum of its parts. Therefore, the parts must be studied and understood within the context of the larger whole. This holistic thinking moves the focus of research from objects to relationships. Studying learning in a classroom setting did not begin until the 1960s, yet this is where the majority of children's academic studies take place. Students are taught within groups of students, yet learning is a very individual activity.

It is important for the teacher to know the history of each student in order to understand what attractors are present and what control parameters might facilitate change. However, this is difficult to do because teachers are involved with groups of students in such a way that time and energy prohibit study of the various systems that make the student who he or she is. So instead, teachers use a variety of learning techniques in order to bring about learning in the most students possible.

Besides the students being made up of different systems, the teacher also is a network of systems with his or her attractor states and state space. This will affect what types of pedagogical practices he or she will use in a classroom setting. For some teachers, teacher-centered strategies are what they are comfortable using, and despite information to the contrary, they will continue in that attractor state because of its stability.

Jane L. Abraham

See also Continuity and Discontinuity in Learning; Piaget's Theory of Cognitive Development

Further Readings

Capra, F. (1996). *The web of life: A new scientific understanding of living systems.* New York: Anchor.

Edelman, G. M. (1992). *Bright air, brilliant fire: On the matter of the mind.* New York: Basic Books.

Gagne, E. D. (1985). *The cognitive psychology of school learning.* Boston: Little, Brown.

Gleick, J. (1987). *Chaos.* New York: Penguin.

Kelso, J. A. S. (1995). *Dynamic patterns: The self-organization of brain and behavior.* Cambridge: MIT Press.

Moore, D. S. (2001). *The dependent gene: The fallacy of "nature vs. nurture."* New York: Henry Holt.

Robertson, R., & Combs, A. (Eds.). (1995). *Chaos theory in psychology and the life sciences.* Mahwah, NJ: Lawrence Erlbaum.

Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action.* Cambridge: MIT Press.

Thelen, E., & Smith, L. B. (1998). Dynamic systems theories. In R. M. Lerner (Ed.), *Handbook of child psychology: Vol. 1. Theoretical models of human development* (5th ed., pp. 563–634). New York: Wiley.

von Glaserfeld, E. (1984). An introduction to radical constructivism. In P. Watzlawick (Ed.), *The invented reality* (pp. 17–40). New York: Norton.

DYSLEXIA

Dyslexia is one of several types of learning disabilities that occurs in children who, despite strengths in other academic and cognitive abilities, experience extreme difficulty in learning to decode and spell printed words. This problem at the level of the single word impedes their ability to fluently read and comprehend connected text. Importantly, their reading difficulties do not stem from lack of educational opportunity, sensory acuity deficits, or socioeconomic disadvantage, although educational opportunity and socioeconomic status may interact with dyslexia to either reduce or exacerbate its severity. The International Dyslexia Association defines dyslexia as follows:

> A specific learning disability that is neurological in origin. It is characterized by difficulties with accurate and/or fluent word recognition and by poor spelling and decoding abilities. These difficulties typically result from a deficit in the phonological component of language that is often unexpected in relation to other cognitive abilities and the provision of effective classroom instruction. Secondary consequences may include problems in reading comprehension and reduced reading experience that can impede the growth of vocabulary and background knowledge.

Estimates suggest that dyslexia occurs in 6% to 17% of the school-age population depending on how it is defined. Regardless of prevalence estimates, children with dyslexia comprise the bulk of children receiving special education services in the United States for students with learning disabilities. Although

dyslexia manifests in childhood with the onset of formal reading instruction, it has definable precursors that permit early detection before formal schooling. In the absence of targeted intervention, the deficiencies observed in childhood continue to be evident in the same individuals well into adulthood. Dyslexia may also occur concomitantly with other conditions, including other learning disabilities involving mathematics (dyscalculia) and written expression (dysgraphia), oral language disorders, and attention deficit disorder with or without hyperactivity. These co-occurrences are usually comorbid, meaning that the child has more than one problem.

Scientific understanding of dyslexia has dramatically increased in the past four decades. It is now widely accepted that dyslexia occurs on a continuum in much the same way as obesity or high blood pressure. What this means in practice is that dyslexia is not an all-or-nothing categorical disability. Instead, it occurs in degrees of severity. Establishing the cut point on the continuum at which word reading difficulties constitute a disability is not well established and accounts for the variability in prevalence estimates. Current classification schemes vary by state and by program type. In the United States, children can receive services for dyslexia as part of special education under the Learning Disabilities (LD) category in the Individuals with Disabilities in Education Act (IDEA), where problems with basic reading skills are one of the eight domains in which LD can occur. A few states, such as Texas and Louisiana, also provide dyslexia services outside of special education. In most states, qualification for dyslexia services under the LD category would be based on an achievement/IQ test discrepancy, with a number of discrepancy models in use from state to state. In the recent reauthorization of IDEA, states were also allowed to use response-to-intervention (RTI) as a mechanism for identifying learning disabilities including dyslexia as well as models based on discrepancies in achievement relative to age. Regulations permitting the use of RTI for identification are just emerging from different states.

Historical Roots and Theories

Over the years, a number of other theories and resulting treatments have gained and lost prominence in the field of dyslexia. A few of the most prominent theories are briefly reviewed below.

Deficits in Visual Perception

Dyslexia was first described more than 100 years ago. By 1900, the condition was referred to as *congenital word blindness* and was believed to be caused by difficulty in storing visual impressions of words. Building on the assumption that the underlying causes of observed reading difficulties were based in deficits in visual processing, Samuel Orton outlined the first fully developed theory of dyslexia, which he called *strephosymbolia* (twisted symbols). His theory suggested that children were experiencing difficulty in establishing hemispheric dominance within the brain, resulting in an inability to suppress mirror image counterparts, which was presumed to cause optic reversibility in visual perception that resulted in seeing symbols as reversed images (e.g., *b* as *d*; or *was* as *saw*).

Although the basic tenets of Orton's theory have not proved to be correct, they have been hugely influential in driving the focus of subsequent theories and treatments of dyslexia even to the present time. It is from Orton's theory of dyslexia that the widely held belief that the treatment of dyslexia requires multisensory methods of instruction stems. The underlying assumption being that by activating the auditory, tactile, and kinesthetic modalities of learning, students are able to compensate for inherent weakness in the visual domain. In fact, the Orton–Gillingham method for teaching reading, or one of many derivatives of this method, is pervasive in the treatment of dyslexia even today. However, empirical support for these methods is mixed. Likewise, there is no indication that the multisensory aspects of this instruction are actually the salient aspects of the instruction. The content of these programs focuses on teaching alphabetic decoding, and this focus likely accounts for any effects achieved. Interventions that do not actually include multisensory aspects but do carefully teach alphabetic decoding have also achieved very positive results with this population.

The view that dyslexia was caused by dysfunction in the visual perception system dominated the field until it was systematically evaluated in the 1970s and 1980s. Outcomes from these studies clearly demonstrated that the visual perception ability of dyslexic and normal readers was not different, and that visual perception ability does not predict performance of reading ability.

Low-Level Visual Deficits

Dyslexia has also been attributed to visual tracking problems assumed to be caused by ocular motor

deficiencies; visual masking effects associated with deficits in the transient visual system; and scotopic sensitivity syndrome (SSS) involving excess sensitivity of the retina to particular frequencies of light, causing the brain to distort information. Of these theories, SSS currently has the most prominence in practice. The interventions recommended for SSS involve placing colored overlays over text or using tinted lenses designed to selectively filter out problem frequencies and make the text visually stable.

The visual tracking theory of dyslexia has long been discredited by well-controlled eye movement studies finding no differences between poor and normal readers. Similarly, the evidence to support the theory of visual masking effects has been weakened by the finding that significant numbers of normal readers have similar transient system deficits. Likewise, independent reviews of the results of interventions based on SSS theory have not found specific associations of reading, dyslexia, and other disabilities.

Perceptual-Motor and Cerebellar Deficits

In the 1960s and 1970s, theories commonly attributed the cause of dyslexia to perceptual-motor deficits. Although the specifics of each theory differed, they shared the common belief that a child must have adequate visual-motor functioning as a prerequisite for academic learning, and that motorical treatment was needed to reorganize neurological functioning to allow adequate visual-perceptual ability. The treatments aligned with these theories included such things as having children practice creeping and crawling, walking on balance beams, standing on balance boards, and completing eye-tracking exercises. Although these theories and their accompanying treatments were very popular, they did little to improve the academic outcomes for children and have largely become extinct in practice.

Despite these results, there has recently been some resurgence of this approach to treating dyslexia and other learning disabilities, despite the lack of scientific evidence for effectiveness. This is most apparent in the cerebellar theory of dyslexia proposed by Roderick Nicolson, Angela Fawcett, and Paul Dean. The hypothesis is that children with dyslexia fail to automatize reading and other cognitive and motor skills mediated by the cerebellum. There is little evidence that supports this hypothesis or the recommended interventions.

Low-Level Auditory Deficits

Another theory that has attracted attention more recently is put forth by Paula Tallal, who theorized that children with dyslexia are fundamentally impaired in processing acoustic stimuli characterized by spectral parameters that change rapidly in intensity. Although originally based on children with oral language disorders, Tallal expanded this hypothesis to children with dyslexia, suggesting that dyslexia is caused by lower-level auditory problems interfering with the processing of sounds with rapidly changing properties, which in turn impairs speech perception and thus the acquisition of phonological awareness and phonological decoding. Early empirical evidence provided support for this theory, but more recent investigations have not found specific associations of these perceptual problems with dyslexia. The treatments that have emerged from this theory, which involve using synthetic speech to slow down the acoustic stimuli, have not been found to have a major impact on reading skills in randomized, controlled experiments.

Phonological Coding Deficits

Currently, the theory that dyslexia is the result of weak phonological coding has accumulated substantial and converging scientific evidence to support it. Phonological coding is the ability to use speech codes to represent words and parts of words. Weakness in this domain results in poor phonemic awareness and letter-to-sound mapping, resulting in poorly specified phonological representation of words. These coding deficits result in difficulty in storing and retrieving words as unitized orthographic representations, and difficulty in processing alphabetic information in working memory. The theory suggests that difficulties in storage and retrieval impair the child's ability to form necessary connective bonds between spoken and written words, which in turn impedes the storage of fully specified representations of word spellings and thus impedes fluency of word recognition.

Evidence to support this theory comes from two sources. First, a relation between phonological coding and dyslexia can be inferred because children with dyslexia have consistently been found to perform significantly below normal readers on different assessments of phonological processing, Likewise, a causal link can be inferred because intervention studies directly addressing phonological coding deficits have

repeatedly demonstrated positive outcomes on the reading outcomes of children with dyslexia. There is controversy over the exact manifestations and contributions of different phonological processing domains, but the link of word reading and phonological processing, and the extrapolation to dyslexia, is widely recognized as a major scientific discovery.

Neurobiological Correlates

Dyslexia has always been assumed to be neurobiologically based. However, only recently have the neurological and biological causes of dyslexia been able to be examined directly rather than just inferred.

Genetics

Studies of the heritability of dyslexia show that there is a strong genetic link for reading and reading disability. For example, a child whose parents had reading problems is eight times more likely to experience a reading problem; 25% to 60% of parents of children who are dyslexic also experienced significant reading difficulties; and the rate of concurrent dyslexia among identical twins is above 80%, but below 50% for fraternal twins. The estimates of the genetic contribution ranges from 50% to 80% depending on age, schooling, and other sample characteristics. Furthermore, recent studies have identified a constellation of specific genes involved in dyslexia. Multiple researchers have identified an area on chromosome 6, and chromosome 15 has been identified by several researchers. Potential markers on chromosome 1 and 2 have also been reported but not replicated.

Even though there is a clear genetic link to dyslexia, these studies do not indicate that the genetic links are specific to dyslexia, but instead are linked to reading and the variability in reading skills that characterizes any population. In addition, the same body of research demonstrates that the environment plays an important role in determining reading outcomes for individual children. Whether or not significant reading problems will develop in a child with a genetic predisposition for dyslexia also depends on what types of reading experiences that child is provided in his or her home and school environment. Even if heritability is strong, the genetic contribution does not mean that the reading problem is immutable to intervention, although more intense and targeted instruction may be needed.

Brain Function

The evidence that dyslexia is greatly influenced by environmental factors is supported in recent neuroimaging studies of brain function in students with and without dyslexia. In these studies, various neuroimaging techniques, such as positron emission tomography, functional magnetic resonance imaging, and magnetic source imaging, are used to assess the brain's response to cognitive challenges. The findings from studies using these techniques converge to show that tasks requiring reading are associated with increased activation with the basal surface of the temporal lobe, the posterior position of the superior, the middle temporal gyri, and the inferior frontal lobe. Magnetic source imaging studies, which allow time sequences within the brain to also be mapped, demonstrate that in normal readers, the occipital areas of the brain that support primary visual processing are activated first, followed by regions in the basal temporal areas in both right and left hemispheres. This is followed by simultaneous activation in three areas of the temporal and parietal areas corresponding to the superior temporal gyrus, Wernicke's area, and the angular gyrus, often more predominantly in the left hemisphere. Students with dyslexia activate the same regions of the brain, but there is often much less left hemisphere activation and differences in the temporal sequences of brain activation.

More recently, a series of studies has been conducted to determine the impact of carefully designed instruction on phonological coding on the brain activation patterns of children with or at risk for dyslexia. In each of these studies, the brain activation patterns of children at risk for dyslexia (i.e., 5- and 6-year-olds) or identified as dyslexic (i.e., 7- to 17-year-olds) show predominantly normalizing patterns after well-designed, intense interventions. Importantly, this finding has been found among both young readers prior to the onset of reading difficulties (i.e., intervention in first grade) and older children and adults who have experienced serious reading difficulties. These studies indicate that the neural systems that mediate word recognition and dyslexia are malleable and responsive to intervention.

Effective Treatment

Since the early 1990s, a number of studies have carefully evaluated interventions designed to prevent or

remediate dyslexia. From this body of research, we now know much about what constitutes effective intervention for these children. Perhaps the most important finding is that the needs of children with dyslexia are not unique to only children with dyslexia. The same type of instruction that is effective with these children is also highly effective for other children who experience reading difficulties for other reasons, such as economic disadvantage. Likewise, the critical components of this instruction are the same for even normally developing children, although the level of intensity and amount of repetition needed will be greater for children with dyslexia.

This critical content includes explicit and strategic instruction to ensure that children develop high levels of phonemic awareness; in-depth knowledge of letter-sound correspondence, including the many spelling variations; acquisition of accurate, quick, and flexible word recognition skills moving beyond the single syllable; smooth, fluent oral reading through adequate practice reading increasingly more complex text; and strategies for the deep processing of text. For the treatment of dyslexia, the most crucial component of instruction is the inclusion of explicit instruction in fluent word recognition that includes phonological recoding (i.e., sounding out words) as a word recognition strategy. Phonological recoding appears to account for individual differences in word recognition in both children and adults and is at the heart of phonological coding problems. When instruction is sufficiently intensive and integrates these various components into daily instruction that cumulatively becomes increasingly more complex, children with dyslexia improve considerably in not only accuracy of their word recognition skills, but in their overall reading ability. Particularly striking results are apparent among young children who have not yet experienced reading difficulties, but who have risk characteristics of dyslexia. In multiple studies, the occurrence of reading problems among students at risk for dyslexia and other reading problems has been reduced to below 1.5% of the total population.

One of the most exciting developments in reading research is dramatic growth in researchers' understanding of which skills predict a phonological processing core deficit. Today, researchers can confidently predict which children are most at risk for developing dyslexia and other reading problems as young as kindergarten age. Thus, it is now possible to identify risk for dyslexia early, provide high-quality intervention before failure has occurred, and in most cases normalize reading ability. Early literacy skills that have been demonstrated to accurately predict risk in young children include (a) phonemic awareness; (b) knowledge of letter names and sounds; and (c) speed of lexical retrieval measured through the rapid naming of objects, colors, digits, or letters.

For children who are not identified early, the outlook is not as positive, with reading problems becoming more difficult to remediate by the end of third grade. Older children require more intense interventions, occurring for longer periods of time, than do younger children. Although the impact of early preventive and later remedial approaches are comparable, older children are usually so far behind that it is difficult to provide the intensity and reading experience that will make them completely proficient readers. Although the outcomes for these groups of children can be very impressive, major concerns remain about the development of fluent reading ability among older students even after intervention, partly because of lack of reading experience. Likewise, these children often have persistent deficits in vocabulary and comprehension that also may reflect an earlier inability to access print, as well as co-occurring problems with oral language development that affect listening and reading comprehension independently of phonological processing.

Instructional Programming

Programming to prevent or treat dyslexia and other reading difficulties has proven a great challenge within the public schools. Current proposals recommend a multitiered model of reading intervention. In most of the proposed multitiered models, Tier 1 is improved classroom-level general education core reading instruction so that critical content is taught well from the beginning. Children who experience difficulty in spite of receiving quality classroom-level instruction move into more intense, small group intervention (Tier 2) provided in addition to their core (Tier 1) instruction. The focus is not on labeling children, but rather on providing services to children at risk for reading failure for any reason. Only after a child fails to respond adequately to the first two levels of instruction would a child be considered dyslexic, requiring Tier 3, or ongoing, support. Tier 3 intervention is typically described as having greater intensity and duration than Tier 2 intervention. Different models

vary in the point at which special education eligibility and services would be provided.

Understanding Dyslexia

In the past few decades, much progress has been made in understanding both the causes and treatment of dyslexia. It is now understood that dyslexia is a neurobiologically based disorder of phonological core processing ability that causes individuals to experience severe difficulty in reading and spelling at the level of the single word. Furthermore, it is now known that although genetics plays a large role in determining if a child has a predisposition for developing dyslexia, the environment also plays a large role in determining the level of dyslexia expressed in any one individual. Environments that provide repeated and careful opportunities to practice phonological coding will result in a decreased expression of dyslexia. Furthermore, dyslexia is treatable, even among older children, although outcomes are generally better for younger children. However, this is simply because younger children have not yet fallen behind and thus don't have as much ground to cover in order to catch up.

Patricia G. Mathes and Jack M. Fletcher
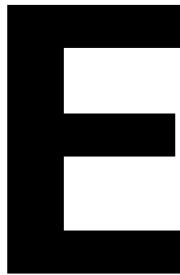
See also Learning; Learning Style; Reading Comprehension Strategies; Special Education

Further Readings

Aylward, E. H., Richards, T. L., Berninger, V. W., Nagy, W. E., Field, K. M., Grimme, A. C., et al. (2003). Instructional treatment associated with changes in brain activation in children with dyslexia. *Neurology*, *22*, 212–219.

Eden, G. F., Jones, K. M., Cappell, K., Gareau, L., Wood, F. B., Zeffireo, T. A., et al. (2004). Neural changes following remediation in adult developmental dyslexia. *Neuron*, *44*, 411–422.

Fletcher, J. M., Lyon, G. R., Fuchs, L. S., & Barnes, M. A. (2007). *Learning disabilities: From identification to intervention.* New York: Guilford.

Mathes, P. G., Denton, C. A., Fletcher, J. M., Anthony, J. L., Francis, D. J., & Schatschneider, C. (2005). The effects of theoretically different instruction and student characteristics on the skills of struggling readers. *Reading Research Quarterly*, *40*(2), 148–182.

Orton, S. (1928). Specific reading disability-strephosymbolia. *Journal of the American Medical Association*, *90*, 1095–1099.

Pakorni, J. L., Worthington, C. K., & Jamison, P. J. (2004). Phonological awareness intervention: Comparison of Fast For Word, Earobics, and LiPS. *Journal of Educational Research*, *97*(3), 147–157.

Schatschneider, C., Francis, D. J., Carlson, C. D., Fletcher, J. M., & Foorman, B. R. (2004). Kindergarten prediction of reading skills: A longitudinal comparative analysis. *Journal of Educational Psychology*, *96*(2), 265–282.

Shaywitz, B. A., Shaywitz, S. E., Blachman, B. A., Pugh, K. R., Fulbright, R. K., Skudlarski, P., et al. (2004). Development of left occipito-temporal systems for skilled reading in children after a phonologically-based intervention. *Biological Psychiatry*, *55*, 926–933.

Simos, P. G., Fletcher, J. M., Sarkari, S., Billingsley-Marshall, R. L., Denton, C. A., & Papanicolaou, A. C. (2007). Intensive instruction affects brain magnetic activity associated with oral word reading in children with persistent reading disabilities. *Journal of Learning Disabilities*, *40*(1), 37–48.

Stanly, G., Smith, G. A., & Howell, E. A. (1983). Eye movements and sequential tracking in dyslexic and control children. *British Journal of Psychology*, *74*, 181–187.

Torgesen, J. K., Wagner, R. K., Rashotte, C. K., Rose, E., Lindamood, P., Conway, T., & Garvan, C. (1999). Preventing reading failure in young children with phonological processing disabilities: Group and individual response to instruction. *Journal of Educational Psychology*, *91*, 579–593.

Vaughn, S. R., Wanzek, J., Woodruff, A. L., & Linan-Thompson, S. (in press). A three-tier model for preventing reading difficulties and early identification of students with reading disabilities. In D. H. Haager, S. Vaughn, & J. K. Klingner (Eds.), *Validated reading practices for three tiers of intervention.* Cambridge: MIT Press.

Vellutino, F. R., Fletcher, J. M., Scanlon, D. M., & Snowling, M. J. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychiatry and Psychology*, *45*, 2–40.

Web Sites

The International Dyslexia Association: http://www.interdys.org

E

A master can tell you what he expects of you. A teacher, though, awakens your own expectations.

—Patricia Neal

Early Child Care and Education

Early child care and education is defined as any care on a regular basis by someone other than a child's immediate family members. Infant and toddler care refers to this type of care for children from birth to age 3 years. Early care and education are important to educational psychology because they are associated with later cognitive and socioemotional outcomes for children in educational settings. In addition, with the historical changes that resulted in parents working outside of the home, increases in wraparound child care to include hours before and/or after the school day have become normative for school-age children, with implications for educational psychology. This review first defines the topic of early child care and education and outlines the different settings in which care takes place. This is followed by an explanation of the importance of early care and education environments to child development outcomes. Finally, the effect of family contexts in early care and education as they relate to child development outcomes is presented.

Types of Early Child Care and Education

Three main types of early care and education exist for infant and toddler groups. These are in-home care, services in child care homes, and services in child care centers. *In-home care* refers to care when an individual comes into the family's home to provide care. *Child care homes* and *centers* are out-of-home services where children receive care in another individual's home or at a location other than someone's home, such as a child care facility.

Importance of Early Child Care and Education Environments

The impact of physical environments and subsequent early experiences on early brain development makes consideration of early care and education environments important. Given the primary importance of the family to early child development, child care has been found to make a unique contribution in the area of cognition. This unique contribution is moderated by three known variables: the amount of time children spend in care, the quality of care, and the type of care. Specifically, the incidence of problem behavior and minor illness is higher in children who spend more than 30 hours per week in care. Regarding quality, infants and toddlers involved in high-quality child care have been shown to have better social cognition and better reading and math skills compared with children in low-quality settings. Last, regarding type of care, center-based care, compared with other types of care, is associated with better language, social skills, and pre-academic skills with respect to letters and numbers.

Models

Two important models of early child care and education exist that encompass both regulatory and process benchmarks for early care and education. The first is the Early Head Start program, administered by the U.S. Department of Health and Human Services, Administration for Children and Families, which is aimed at providing services to families with infants or toddlers living below the federal poverty level. The second is the military model of child care, instituted by the U.S. Department of Defense, aimed at providing early care and education services to military families.

These models share the philosophy that incorporating both child- and family-centered practices is integral to positive child outcomes. It is clear that the incorporation of both child- and family-centered practices is important from evaluation of the effects of child care in the lives of infants and toddlers from these populations. Specifically, it is clear from evaluation of the Early Head Start model, in particular, that over time children do better when parents have higher levels of education and income, when mothers have fewer symptoms of depression, and when families incorporate learning activities into their home, such as books and play materials, and are involved in their young children's play activities.

The Early Head Start Program Model

The Early Head Start program model is built on a framework encompassing four cornerstones to ensure best practices in delivering services to infants and toddlers and their families. These cornerstones include child development, family development, community building, and staff development. Within the cornerstone of child development, which includes child care services, program design incorporates many objectives related to quality child care. Specifically incorporated are the objectives of positive child health and development, education and early childhood development, child health and safety, good nutrition, and good child mental health. Within the objective to promote positive child health and development is the identification of activities that will foster children's physical, cognitive, and socioemotional growth in a child care setting. Because of the infant/toddler age of the children enrolled in these programs, emotionally secure parent–child as well as child care provider–child

relationships are seen as an integral component of fostering these areas of development. Thus, in addition to developmentally appropriate early care and education services, parental supports are in place to build strong parent–infant emotional connections. Supports include activities such as center- or home-based parent education as well as parental involvement in the child care program.

Community building is a component of the Early Head Start framework that stems from ecological theory, which recognizes the important interaction between the individual and larger systems. For families with infants or toddlers and limited resources, the Early Head Start programs have aimed to increase access to resources in the community. The comprehensive nature of the program has been shown to be important to families with limited resources who have infants or toddlers.

Staff development is a last important cornerstone to the Early Head Start programs that is universally important to child care. Teachers with more advanced training and education are also more likely to understand and use developmentally appropriate practices in the classroom. In fact, it has been established that developmentally appropriate behaviors in child care staff will increase directly as the number of hours of training received increases. Training, supervision, and mentoring in understanding and developing child socioemotional development are integral activities implemented in the Early Head Start program model to achieve the goal of staff development for the larger purpose of enhancing positive child outcomes.

The Military Model of Child Care

The U.S. military has developed a model of child care that exemplifies the way intensive reform can drastically improve a service such as child care. Before the Military Child Care Act of 1989, the military model of child care was very similar to, and plagued by, the same struggles as current civilian child care. The military model of child care came about as a result of a demographic shift in military personnel in the 1980s, as more men and women with families chose the military as a career. The low-quality child care arrangements available to military families before 1989, such as unsafe conditions, poor and unenforced standards, and extremely high staff turnover rates, hampered military readiness. For the military, such issues implied the need for an

improved child care system. This need was addressed and accomplished by creating a systematic approach to improving its child care system.

Today, the U.S. military has set in place a coordinated system for all of its branches. The key components of the system aim to improve quality through enforced standards, accredited programs, well-trained and well-compensated staff, and an increase in family-identified goals, such as affordability. The military model of child care systematically addresses universal needs for child care.

The military model of child care has improved its program quality by creating firm standards in all areas of care, including health, safety, staff–child ratios, and staff training. These standards exceed some state standards. Certification is done annually, and regulations are enforced through quarterly, unannounced inspections of child care programs. Meaningful sanctions are in place for when programs are found to be noncompliant with standards. Inspectors are given the latitude to deem a program necessary to be fixed, waived, or closed, reinforcing the intolerance for unsafe programs to remain open. The civilian child care community would benefit greatly from strict enforcement of quality standards as exemplified in the military model.

The military model incorporates a single point of entry approach to child care. This addresses the differing needs and wants of families as it pertains to care for their child(ren). Child care centers, family child care homes, and individual providers are maintained in one, current database that makes it easier for families to choose what is best for their family. This single point of entry method has two additional, positive consequences. First, it is easier for accreditation and monitoring to take place within a universal system. Second, resource and referral services are important to families choosing child care, and a universal system can be cost-effective to this important first step to finding child care for families. In civilian child care, some states have incorporated resource and referral services, yet many families are not aware of this service. Public information campaigns address this issue. Additionally, a statewide resource and referral system that links program accreditation and monitoring is a cost-effective way to address many issues simultaneously, such as family need for child care availability information and tracking of program accreditation and monitoring.

The National Association for the Education of Young Children (NAEYC) is an agency that rewards child care programs for going beyond mandatory licensing requirements through an accreditation system. In the military system, NAEYC provides financial assistance to child care programs. This accreditation standard in the military system has led 99% of centers to be accredited.

Moreover, half of children in military child care are under 3 years old. As a result of the large number of infants and toddlers in child care, improvement for this vulnerable population has been a priority in military reform. Within this system, children are able to enter at 6 weeks old and for up to 12 hours per day, as this is what is needed for many military families. An additional need, for many military families, is child care services provided during "off" hours, such as during evenings or on weekends. As the economic environment changes, civilian families face circumstances similar to those of military families; for example, extended family may not be in the same vicinity. Thus, the need to rely on professional support for child care is a reality for both military and civilian families. The military model of child care has focused reform on providing flexible care schedules to families.

To address the needs of infants and toddlers, military child care policy has also instituted primary caregiver assignments of one child care provider to a child. Additionally, infants are allowed individual sleeping and feeding schedules and sleep in cribs with the rest of the group, rather than in a separate sleeping room. Policy has also emphasized safety and sanitation for this age group. Child care centers have incorporated building design that decreases the spread of germs and reduces child accident rates through exitways from infant and toddler classrooms to outdoor play areas designated specifically for different age groups.

The military model of child care incorporates policies above and beyond state licensing requirements. Policies such as standard use of surveillance cameras in the classrooms and a diaper-changing rotation among staff (so that no one individual is responsible for diapering) can decrease the incidence of child abuse.

Last, within the military model, parents are seen as constituents and are, thus, involved in programs—for example, through their presence in the centers and through their decision-making capacity. Programs are designed to be family-friendly and support infant and toddler health outcomes by providing rooms for mothers to nurse their infants and toddlers. Parents also

participate in advisory councils as well as serve on inspection teams to monitor child care quality standards.

The military model of child care emphasizes staff training and compensation. The system requires that all employees receive basic preservice training. Additionally, child care personnel begin with higher-than-average wages, and they are also rewarded with increased wages for achieving training milestones.

State Licensing, Monitoring, and Optional Accreditation

In general, states regulate early child care and education through licensing standards and monitoring practices. Many programs choose not only to incorporate standards that meet the required state licensing standards but also to adopt standards above and beyond minimal state standards. One organization, the NAEYC, is widely used by child care programs and is the largest organization serving on behalf of infants and children through age 8 in child care. The network was founded in 1926 with the mission of helping programs optimize child outcomes and to help families identify high-quality child care programs. It has an alliance of local, regional, national, and international affiliates. To be accredited by NAEYC, a program must involve itself in (a) self-study, through which staff identify areas for improvement; (b) validation, in which program information is verified during a site visit by a team of trained volunteers; and (c) a review by recognized experts who judge whether the program is in compliance with the accreditation criteria. Programs are granted accreditation for 3 years with agreement to improve areas in which they are only marginally compliant. Annual reports are required to document continued compliance and improvement.

Regulated and Process Variables in Early Child Care and Education

Both the U.S. Department of Defense and the Early Head Start programs are large-scale models with systematically incorporated standards above and beyond many state standards for infant and toddler early care and education. Many of the standards associated with these two programs come from NAEYC, the National Association for Family Child Care, or both.

Two types of variables are discussed in relation to early child care and education monitoring and evaluation: *regulated* variables and *process* variables.

Regulated variables encompass health and safety system variables, including adult-to-child ratios, group size, and staff education levels. Regulated variables matter to infant and toddler care in that they uniquely predict child development outcomes separate from other variables. Process variables relate to developmental outcomes in early care and education, both cognitive and socioemotional, which come from the relationship between providers of early care and education and the children in their care.

Regulated variables are predictive of process outcomes, which are related to child development outcomes. Thus, early care and education programs that have smaller adult-to-child ratios, smaller group sizes, and providers with higher education levels predict more positive provider-to-child social interactions. This, in turn, positively affects children's cognitive and socioemotional development. Thus, the most effective early care and education programs incorporate both highly regulated and process variables into services for infants and toddlers.

Family Contexts in Early Child Care and Education

Family features have been found to be the most predictive determinants to child development outcomes for children in early care and education settings. From these findings, models of early care and education have been established that emphasize the components of positive family features that can be replicated in early care and education settings. These process variables are associated with cognitive and socioemotional development by building secure relationships between early care and education providers and the infants and toddlers in their care. They include discreet aspects of emotional connections between care providers and infants and toddlers.

To develop emotional connections between care providers and the children in their care, the physical as well as interpersonal environment needs to be considered. Regarding the physical environment, standards call for a more home-like than school-like environment for infants and toddlers. To meet the needs of infants and toddlers, physical environments that replicate home environments often include primary caregiver assignments of one child care provider to a child. Additionally, infants are often allowed individual feeding and sleeping schedules and sleep in cribs among the rest of the group, rather than in separate sleeping

rooms. Children's work is displayed in the classroom and at child eye level to build a sense of ownership and belonging. Equipment, such as tables and chairs, is child-sized, and soft surfaces exist for children to rest and relax on during the day. A variety of materials exist, especially dramatic play materials, for toddler-age children to foster socioemotional and social skills development through role-playing behavior. Also, providers interact with the physical environment in ways that promote prosocial behavior, such as engaging in dramatic play activities and modeling and facilitating prosocial behavior between children. Last, a predictable environment with few changes to the setting is supportive of a child's emotional security.

In addition to the physical environment, socioemotional development in early care and education is also supported by way of emotional connections through caregiver–child interactions. An essential aspect of preparing child care providers to foster socioemotional development in the child is to learn about how to form emotionally available and securely attached relationships.

Attachment and Emotional Availability in Relationships in Child Care

Traditionally, attachment is viewed as a quality of parent–child relationships; however, children may also develop attachments with adult caregivers other than parents. Whereas approximately 70% of children form secure attachments with their parents, estimates indicate less than half of children in child care form secure attachments with their child care providers. The presence of caregiver sensitivity, appropriate structuring, nonintrusiveness, and nonhostility, which are qualities of emotional availability, are important, and research indicates that these qualities are predictive of secure attachments and children's emotional availability. According to Zeynep Biringen, caregiver sensitivity in the context of early child care and education includes both nonverbal and verbal behaviors. Nonverbal behaviors include physical affection, affect, awareness of timing, creativity during play, prompt and accurate responsiveness, and respectful behavior. Verbal behaviors include the provider mirroring what the child is doing, singing, and speaking to the child respectfully. Caregiver structuring behaviors that are nonverbal include not restricting the child in equipment such as a high chair longer than needed (or to the point where the child's patience runs out) and setting up individual and group activities. Verbal behaviors include talking to the child in passing, talking during meals and snacks, talking when picking up the infant to take to an activity (diapering), and verbally framing the transition from one activity to another. Caregiver nonintrusiveness behaviors that are nonverbal include picking up a child in a gentle way, limiting the number of physical interventions, reserving abrupt behavior for emergencies, not introducing all of the play toys, and playing with the child by focusing on an object that interests the child. Verbal behaviors include using language that is not intrusive, using a gentle and calming voice, and reserving language that can be jarring for emergencies. Caregiver nonhostile, nonverbal behaviors include attempts to be engaged with children rather than showing boredom and using coping mechanisms to limit stress. Verbal behaviors include limiting negative talk and refraining from complaining. Last, caregiver awareness of the child's contribution to interactions is evidenced by caregiver behaviors such as realizing that the child's emotions are linked with child–provider interactions; showing awareness that a child who is bland in facial expressions but causes no trouble and is undemanding is not a preferable presentation; and knowing that when a child is acting out, it is a communication mechanism. These nonverbal and verbal caregiver behaviors discreetly influence child socioemotional outcomes for infants and toddlers, specifically secure attachments, in the early care and education setting. Children with secure attachments are more likely to experience positive long-term outcomes throughout their lives. For that reason, children need caregivers who are emotionally available to them. Some of the benefits children may experience from emotionally available relationships are secure attachment, less aggressive behavior, less victimization, better peer relationships, more attentiveness at school, and better relationships with teachers. Early care providers must take into account the emotional and relational needs of infants and toddlers and facilitate emotionally available and secure relationships.

Despite the importance of the early child care context, new research findings from the National Institute of Child Health and Human Development (NICHD) Study of Early Child and Youth Development indicates only a small-magnitude linkage between child care and child development and that the family is more crucial than early care professionals in a child's development. Because child care has become an important resource and a way of life for working

mothers and fathers, it was important to conduct a definitive study to understand the impact of child care on children. For 15 years, a multisite investigation at 10 U.S. sites were examined, following the development of 1,000 healthy children and their families, beginning with a child's birth. Ethnically diverse families, with some 80% two-parent households, were initiated into the NICHD project. These children were in child care for close to 30 hours per week from birth to almost school age, thus including a sample of participants who had a high dose of child care in their infancy and preschool years. Most began in home-based situations with relatives or nonrelative caregivers and then progressed to center-based facilities.

An important finding from this multisite investigation is that children do better when parents are more educated; when income for the family is higher; when mothers are low on symptoms for depression; and when families have routines, rituals, books, and other learning and play materials. The children in the study had better outcomes if mothers were sensitive, and mothers were more likely to be so if they had more socioeconomically advantaged lives. These parent or family qualities were up to three times more strongly associated with child development than child care features, although when child care features were examined, qualities such as higher-quality care (with small child-to-provider ratios, good play materials, etc.) were found to be important. Those children who experienced high-quality care were more likely to have better social and cognitive (reading, math) skills than those who experienced low-quality care.

To bring a different perspective into focus, research on attachment has indicated that sensitive mothers are sometimes in relationship with insecure infants and toddlers. In such mismatched situations, "marginal" child care has been found to be a part of the equation. In a study conducted in Israel using the full spectrum of socioeconomic levels, infant attachment was assessed using the well-known Strange Situation procedure, and maternal sensitivity was assessed using the Emotional Availability Scales. Results showed that the hypothesized link between infant secure attachment and maternal sensitivity was found only for infants in individual care at home and not for infants at centers of low quality. It is possible to hypothesize that extensive experience in an early care environment that is not in tune with an infant's needs and that misses important cues for interaction and relationship may be akin to small doses of trauma. When an infant spends whole days where he or she is not responded to in an appropriate fashion and where engagement is lacking (on both the side of the professional and the side of the infant)—that is, where there is no real relationship—the infant may take back into the home setting behaviors that create insecure attachments to the primary caregivers as well. Thus, for children in child care, it is crucial to focus on both family and child care providers as agents of socialization.

In summary, emphasis on child care research should be to include parents as the primary socialization agents in their children's lives. Intervention science, however, should not abandon child care professionals as agents of socialization.

Shannon Altenhofen,
Brooke Davy, and Zeynep Biringen

See also Attachment; Emotional Development; Family Influences; Head Start; Parenting Styles; Social Development

Further Readings

Aviezer, O., & Sagi-Schwartz, A. (2003). Ecological constraints on the formation of infant-mother attachment relations: When maternal sensitivity becomes ineffective. *Infant Behavior & Development*, *26*, 285–299.

Biringen, Z. (2000). Emotional availability: Conceptualization and research findings. *American Journal of Orthopsychiatry*, *70*, 104–114.

Bredekamp, S., & Copple, C. (Eds.). (1997). *Developmentally appropriate practice in early childhood programs* (Rev. ed.). Washington, DC: NAEYC.

Howes, C., Galinsky, E., & Kontos, S. (1998). Child care caregiver sensitivity and attachment. *Social Development*, *7*(1), 25–36.

Howes, C., Whitebook, M., & Philips, D. A. (1992). Teacher characteristics and effective teaching in child care: Findings from the National Child Care Staffing Study. *Child and Youth Care Forum*, *21*, 399–414.

National Institute of Child Health and Human Development. (2006). *The NICHD Study of Early Child Care and Youth Development (SECCYD): Findings for children up to age 4½ years* (NIH Publication No. 05-4318). Washington, DC: U.S. Government Printing Office.

Web Sites

Early Head Start National Resource Center (EHS NRC): http://www.ehsnrc.org

National Association for the Education of Young Children (NAEYC): http://www.naeyc.org

Zero to Three: http://www.zerotothree.org

Early Intervention Programs

The term *early intervention (EI) programs* refers to a mosaic of services and programs created for young children and their families to meet educational, social, and health needs and promote enhanced learning, social, and communication outcomes. EI programs are provided before a child starts school, sometimes beginning before birth and continuing to school age. A goal of EI programs is to reduce gaps in child development that are due to one or more of the following: poverty or welfare status, family violence, child abuse, lack of family housing or ability to parent, poor health, single parenthood, teen parenthood, low maternal education, lack of proficiency in the English language, disabilities, developmental delays, poor language or emergent literacy, and the presence of a variety of at-risk conditions. EI programs incorporate a range of approaches, including child-focused interventions, dyadic interventions between an infant and parent, and family interventions that include siblings, parents, and even grandparents. The Division of Early Childhood for the Council of Exceptional Children identified a list of 1,606 recommended EI practices based upon a review of empirical research in 48 journals of effective early intervention programs or approaches.

Over the past 35 years, three major areas of interest have evolved in the field of EI programs:

1. Longitudinal experimental research of the efficacy of programs for children from birth to 3 years and preschool age

2. The national establishment of programs to achieve "equity" in education for children in poverty (e.g., Early Head Start and Head Start) and children with special health care needs or disabilities (e.g., Early Intervention Programs for Infants and Toddlers with Disabilities under Part C of the Individuals with Disabilities Education Act [IDEA] and Preschool programs under Section 619 of Part B, IDEA)

3. Concerns for quality and accountability prompted by legislative mandates

This entry will briefly highlight recent research in each of these interest areas and provide additional background information on modes of service delivery in health settings, homes, and centers and the economic advantages provided by intervention services.

Prevention Efforts

Early intervention programs may have features that are universal or are intended for a broad population. All young children are in need of a medical home that will provide continuity of medical care with developmental screening and preventive health care in a family-centered context. Model programs such as Healthy Steps, a pediatric care intervention, provide anticipatory guidance through child developmental surveillance, promote positive health practices, and encourage parents to ask questions and express any concerns they may have regarding their child's development. This program has proven successful in increasing a child's access to well-baby care, because participating families tend to remain with a practice for longer periods, and in reducing negative parenting practices and allowing depressed mothers a safe environment to discuss their distress.

Prevention approaches can begin even before the birth of a child in the form of preventing or ameliorating the effects of conditions that may be caused by environmental, lifestyle choices, or other addressable conditions. A number of early intervention programs have been developed that target women in the neonatal period, particularly women with conditions that require extended hospitalizations. The Newborn Individualized Developmental Care and Assessment Program (NIDCAP), developed by Heidelese Als and others, has fostered new approaches to neonatal care, especially among premature infants. The NIDCAP approach trains neonatal staff to support premature infant regulation and appropriate development through individualized observation. Because of recent technological advances, very low birth weight (VLBW) premature infants have an improved survival rate but are at higher risk for adverse health problems and developmental delays; neonatal intervention approaches like NIDCAP address these issues. As the rate of premature births continues to climb, the development of effective interventions for the medically fragile premature infant is of increasing concern.

Longitudinal Research Studies in Early Intervention

The U.S. Department of Education's Institute for Educational Sciences promotes experimental, randomized, clinical trials as the desirable method to determine efficacy in education. EI research has benefited from a history of longitudinal studies in which the infants, toddlers, and preschoolers receiving intervention services are now adults. The Carolina Abecedarian Project is a classic example of an EI study that demonstrates high methodological rigor, random assignment to intervention and control groups, and low attrition rates. This study of a model program, spanning 8 years, was designed to close the academic gap between young children from poor families and their more economically advantaged peers. In the first phase, full-day center-based services were provided to the intervention group from infancy to kindergarten entry, augmented by a developmental curriculum and an emphasis on language acquisition. From kindergarten to second grade, a home resource teacher supported and encouraged parents to work with their children at home. Participants continued to show consistent cognitive and academic gains compared with controls when assessed at 12-to-54 months and 12, 15, and 21 years of age. Participants in the preschool intervention had fewer retentions, were less likely to require special education services, and, by age 21, had completed significantly more years of education than controls.

Long-term follow-up programs contribute substantially to the field of early intervention, providing essential information on the types of developmental outcomes resulting from the application of models of intervention that vary based on format, content, intensity, and population served. Programs such as Chicago Child-Parent Centers and the Perry Preschool Project have yielded positive long-term outcomes through a combination of center-based and home-based parent-directed programs. In the case of the Perry Preschool Project, although significant gains on cognitive measures were not retained over time, real-life benefits included higher graduation rates and college attendance, improved rates of gainful employment, and reduced crime and teen pregnancy. Studies of specific populations, such as children exposed to cocaine, provide additional support for the beneficial effects of center-based programs in combination with home visits. A new area of promising research is the state implementation of universal prekindergarten programs for 4-year-olds, such as the efforts underway in Oklahoma, which provide encouraging findings of enhanced school readiness with significant gains in areas of math and language achievement. Finally, any mention of longitudinal intervention efforts must include a discussion of Head Start, the most widely offered early childhood program for preschool-age children living in poverty. Given the breadth and impact of Head Start and its companion program for children from birth to age 3, Early Head Start, each of these programs will be discussed in more detail in later sections of this entry.

Looking toward the future, the Early Childhood Longitudinal Studies-Birth Cohort, a nationally representative sample of 10,688 infants born in 2001, provides a prospective opportunity to study child development from infancy to kindergarten. In addition to comprehensive data about child health, family health and circumstances, and direct measures of a child's mental, motor, and social behavior, information about the provision of early intervention services and day care is also available. Eventually, analyses of these data could provide information that is relevant for national policies on children and EI programs.

Home-Based Programs

Within the field of EI, family-centered services are viewed as an integral component of successful programs. Offering services within a home-based setting has been found to contribute to a family-centered approach, as the child is treated within the family system, and caregivers have the opportunity to view and participate in the intervention activity. However, despite the extensive use of home-based services, little research has been done to evaluate the efficacy of this mode of service delivery for children with developmental disabilities. The field has not yet reached a consensus regarding the effectiveness of a home-based model of services, and some studies question the feasibility of implementing a home-based model within early intervention. In particular, concerns exist regarding the training and background of personnel, large provider caseloads, and the cultural mismatch between providers and participating families.

To date, few studies of home-based services document significant differences in participant versus control outcomes. For example, group differences were not found in home-based interventions aimed at improving parent–child interaction and positive parenting practices,

promoting child development, and preventing neglect and abuse. However, the negative findings must be placed in perspective, because to impact child outcomes over the long term, it is necessary to work directly with parents and primary caregivers to improve their ability to take better care of their children. In this regard, home-based services targeting parents and parents-to-be have proven to be effective. For example, research by Olds and colleagues as part of the Elmira Prenatal/Early Infancy Program provides evidence that home-based interventions lead to improvement in pregnancy outcomes, child health and development, and families' economic self-sufficiency.

In sum, given the inconsistent findings within the literature, the efficacy of home-based services has not yet been confirmed. However, the research has provided valuable information regarding some of the factors that limit program effectiveness, including infrequent home visits (i.e., due to cancellation or missed appointments), high attrition rates, urgent family needs that may interfere with program implementation, low parent engagement, heavy service provider caseloads, inadequate training of service providers, and a cultural mismatch between participants and service providers.

Disability: Birth to Three

In 1986, with the passage of Public Law (P.L.) 99–457, Education of the Handicapped Act (EHA), later renamed Individuals with Disabilities Education Act (IDEA), U.S. states had the option of establishing systems of early intervention services for infants and toddlers (i.e., from birth through age 2) with disabilities. All states and territories participate in this program and are asked to establish a coordinated and comprehensive system of services along with a method to ensure payment for early intervention services. EI services are multidisciplinary and uniquely designed to fully involve the child's family in the decision-making process through the development of the Individual Family Services Plan. The central focus on families is a distinct part of EI services both from a policy perspective and a programmatic stance. Thus, it is no surprise that most families receive services in the home, with speech, occupational, and physical therapies; special instruction; and developmental monitoring cited as the most frequent services delivered. The median amount of time for weekly services scheduled per child during the first 6 months in EI is 1.5 hours.

Children with diagnosed disabilities (such as Down syndrome) are automatically eligible for services, but states have discretion in the adoption of a definition of developmental delay and whether or not to serve children who are considered at risk for developmental delay. Variation in the eligibility criteria employed by individual states has resulted in variability in the percentage of infants and toddlers served in each state. In 2002, the most recently reported national count of children participating in Part C services ranged from 1% of the population to 9.38% with the overall average being 2.2%. Child characteristics are strongly related to the age at which children enter early intervention. Depending on eligibility criteria established within each state, children with specific diagnosed conditions and those considered at risk for developmental delay may automatically qualify for EI services and therefore, on average, are enrolled earlier than children with developmental delays. The low rate of early entry for children with developmental delays, including speech and language disorders, suggests the challenges of early identification of these disabilities.

National Early Intervention Longitudinal Study

The National Early Intervention Longitudinal Study (NEILS) is a longitudinal study of a nationally representative sample of the children and their families participating in early intervention services of Part C of IDEA. Beginning in 1997, data collection occurred at entry to EI through kindergarten; the data consist of family interviews, service record, and service provider, program director, and kindergarten teacher surveys. Recognizing the importance of family outcomes for young children with disabilities, NEILS found that families were highly satisfied with EI services, and nearly all families felt they were equipped to help their child learn and develop and believed they could advocate on their child's behalf. Although most families had positive experiences, ethnic minorities and families with children in poor health were more than twice as likely to have less positive outcomes.

In terms of child outcomes, NEILS found that children in EI vary widely in terms of their age-at-entry, health status, reason for eligibility, and pattern and level of developmental skills. Children identified as at risk and those with diagnosed conditions were found to enter much earlier than children identified with developmental delay. NEILS has shown that EI

entrants vary greatly in their developmental strengths and needs, reinforcing the importance of an individualized approach. As of kindergarten entry, one third of EI participants were no longer involved in special education and were not identified as having a disability. This finding suggests that a sizeable percentage of children enter EI with developmental problems, but given the appropriate support and intervention services, these children may avoid long-term involvement in special education. However, the majority of EI participants continue to require supportive interventions through preschool and/or school-age special education programs.

Disability: Preschool

P.L. 99–457, in 1986, also contained provisions in Section 619 to extend a free appropriate education to children with disabilities from the age of 3 to 5. Preschool-age children with disabilities are given the same services as older children in special education, including an Individualized Education Program. However, eligibility criteria for entrance into preschool special education often differ from the criteria established for EI. A challenge for states is to ensure that children in EI programs who need to continue receiving special services are not lost in the transition from the EI system to the school system. In addition, a major research question for the preschool population relates to the relative benefits of providing services in inclusive environments as opposed to segregated programs set up solely for children with disabilities. The preschool programs for children receiving special education from age 3 to age 5 can be in an array of settings, including public schools, Head Start, and other inclusive environments. In the 2002 child count of the preschool population under Section 619, about 30% of children were in inclusive early childhood settings, 39% in an early childhood or special education setting, 6% at home, 13% part-time in both early childhood and special education settings, 1% in a residential setting, 5% in a separate school, and the rest receiving itinerant services, and reverse mainstreaming.

Pre-Elementary Educational Longitudinal Study (PEELS)

The Pre-Elementary Educational Longitudinal Study (PEELS) was commissioned by the U.S. Department of Education to describe the school experiences, developmental characteristics, and educational programs of 3- to 5-year-old children participating in Section 619 preschool special education services. Preliminary study findings suggest that males are disproportionately represented, whereas the ethnic composition of preschoolers with disabilities tends to closely reflect the racial representation of the general population. Parental education level corresponds to that of the general U.S. population; however, a larger proportion of participants were reportedly living in poverty. Within the PEELS sample, approximately 50% of participants were eligible for services because of speech and language difficulties, with an additional 25% identified with developmental delay. PEELS found that approximately 25% of preschoolers with disabilities were born prematurely. On average, children in PEELS scored within one standard deviation of the mean on standardized measures of school readiness. However, wide variability was found across disability classifications. Children with mental retardation scored lowest on measures of verbal skills, whereas children with autism had more pronounced difficulties on measures of social skills and behavioral problems.

Poverty Interventions

Head Start

The Head Start program, administered through the U.S. Department of Health and Human Services, is a nationwide program targeting low-income families and their preschool-age children. In its four decades of existence, Head Start has provided child development and related services to over 23 million children. It offers a comprehensive array of services, including educational interventions, parental involvement, and medical services such as dental care and nutrition planning. Congress recently commissioned a study through the Department of Health and Human Services to document the population served in Head Start and determine Head Start's impact on school readiness and the conditions under which the program was optimally effective. The study sample consisted of approximately 5,000 children between the ages 3 and 4 from 84 nationally representative grantee or delegate agencies. At the time of service entry, participants were randomly assigned to either a control group (non–Head Start, receiving non–Head Start related services) or a Head Start program through which families and their children would receive Head

Start related services. Data collection started in the fall of 2002, and children were followed through the spring of their first-grade year. Findings reported at the end of the first year of data collection suggest a small-to-moderate impact of Head Start participation for both 3- and 4-year-olds. Across the four content areas examined (i.e., cognitive development, socioemotional development, health, and parenting practices) children with special needs showed improved outcomes in the areas of parent-reported child health status and receipt of dental care.

Early Head Start

Early Head Start is a program for low-income pregnant women and families with infants and toddlers up to the age of 3 years old. The program was designed by the Administration on Children, Youth and Families in 1994, and the first funds were received in 1995 and 1996. Early Head Start serves a wide array of families, operating in 664 communities and serving approximately 55,000 children. The program targets low-income families that include pregnant women, infants, or toddlers and aims to enhance the health and development of participants while strengthening community and family partnerships. In addition, since 1998, 10% of Early Head Start programs' funded enrollment opportunities must be offered to children with disabilities.

Overall findings of the Early Head Start Impact Report suggest that in the areas of language, cognitive, and social development, participants experienced a positive program impact compared with children in the control group. Participants were less likely to have delays in cognitive and language functioning and significantly more likely to receive Part C EI services than were control group children. However, this latter finding is contentious as some studies have shown that certain Early Head Start families were actually less likely to receive Part C services depending on family demographic characteristics, including ethnicity, linguistic background (i.e., non-English-speaking families), socioeconomic status (i.e., very low-income), and lower educational levels. Despite such findings, there is still wide-ranging support for the pivotal role that Early Head Start plays in connecting families with much needed services, such as those provided by Part C. Several studies have reported on Early Head Start service providers' efforts to assist families in learning how to navigate the system in order to access services.

In addition, Early Head Start staff offers guidance and recommendations to families with children with suspected delays or developmental disabilities.

Although much of the research on Early Head Start home-visiting practices has focused on typically developing children, the recent Early Head Start national study documents positive outcomes for young children with special needs. One explanation for the success of home-based services offered through Early Head Start may be tied to service individualization. By following Head Start performance standards that mandate weekly home-visits and bimonthly socialization, Early Head Start home-based services are tailored to meet each family's unique needs. When necessary, home-visitors provide additional comprehensive services and help to coordinate family service referrals. Positive parental outcomes included greater knowledge of child development, less parenting stress, ability of parents to provide a more stimulating environment at home, and more bedtime reading. In addition, families participating in Early Head Start home-based services were more likely to be involved in education and training activities. With respect to child outcomes, children receiving home-based services had better vocabulary development and were more engaged in playing with their parents or primary caregivers. Raikes and colleagues have noted three primary factors related to the success of home-based services on child outcomes: quantity (i.e., number of visits), quality (i.e., quality of engagement) and child focus (i.e., how much time was spent directly in child development-related activities).

When comparing home-based services to other service delivery options, it is important to bear in mind the complexities inherent in this form of service delivery. Vulnerable families living under stressful conditions are more likely to choose a home-based service option. This poses greater challenges for home-visitors, as some families have had prior unpleasant experiences with the service system. Families receiving home-based services are more likely to live in poverty and to consist of working and single parent families; in addition, young parents are overrepresented. Each of these factors potentially limits the time and effort that families are able to devote to early intervention activities. One study examining the impact of demographic risk factors on Early Head Start outcomes found that the program was less effective for families with more than three risk factors and that those families with the highest number of risk factors often receive home-based or mixed-approach services.

Standards for Evaluating Effectiveness

New federal and state accountability initiatives require that program expenditures be justified based on student educational gains. Early intervention programs, including those administered through Head Start, Maternal and Child Health, and the Department of Education, are not exempt from this scrutiny. Both Part C and Section 619 services of IDEA failed to demonstrate results under these Government Performance and Results Act requirements and were required to develop new outcome measures. The Early Childhood Outcomes (ECO) Center has assisted the U.S. Department of Education in adopting a functional outcome approach that will ultimately yield outcome indicators on three functional skill areas for each state over time. Developing valid and reliable outcome measures is a challenging yet fundamental task necessary for the accurate appraisal of program results. The ECO notes that misuses of data can result when overly narrow outcome measures are chosen, when additional outside information is not taken into account, or when data are misinterpreted.

Another way to judge the effectiveness of a program is to consider its costs relative to its benefits. Evidence is increasingly available supporting the cost-effectiveness of EI programs, highlighting their direct and indirect benefits to society. Some of the longitudinal programs that have been found to be cost-effective include Perry Preschool, the Abecedarian Project, the Chicago Child-Parent Center, and the Elmira Prenatal/Early Infancy Program. The returns from these programs range from $3 to $17 for every dollar invested. In addition, children in intervention groups were more likely to continue in regular education classrooms rather than special education, finish high school, and be gainfully employed.

Implications

To better understand the implications of successful EI longitudinal studies, it would be helpful to examine those interventions that have not produced the intended results. For example, the Early Head Start studies report that infants in families who benefit the most are those with no more than three risk factors (the neediest families would be among those with three to five risk factors). Shonkoff and others note that EI programs may need to be targeted more for children and families with complex social needs, including child abuse. This same reasoning can perhaps be applied to the very high risk populations being served by Early Head Start.

The Infant Health and Development Program, an experimental study of premature infants, offered well-structured center-based interventions that did not yield the expected results. When infants were divided into high and low birth weight groups, more robust gains were found in the cognitive outcomes of those infants considered low birth weight compared with the infants who were considered very low birth weight. These findings are of concern, as the early fragile medical status of very low birth weight infants puts them at high risk for long-term disabilities. Successful intervention efforts may need to be more intensive and individualized to positively impact this group. In addition to taking a more individualized approach to interventions, current best practice within EI includes consideration of broader issues, such as addressing the needs of the family within the context of the community.

To be effective, EI programs must be designed to address the needs of the population served. Three current trends impacting intervention services are the increasing rates of preterm births, child poverty, and a broadening of cultural diversity within the United States. To respond to these needs, broad-based programs are required to promote health prevention efforts focusing on early identification and intervention targeting prematurity and other risky health conditions. All EI programs are faced with changing cultural groups and the responsibility to tailor interventions to the needs of diverse linguistic, cultural, and sometimes national backgrounds. In European countries, tax structures and transfers have been shown to lower child rates of poverty; however, within the United States, poverty rates still exceed 20%, with increases noted in 2002 following a 9-year period of declining poverty rates.

EI programs have been built on the success of well-designed longitudinal studies. Going forward, all federal programs, including Head Start, Early Head Start, and Part C and Section 619 of IDEA, will be expected to produce clear developmental outcomes among children and families served in order to meet accountability standards. The complexities of child development and family issues may be difficult to capture in traditional government accountability processes. In addition, based upon the longitudinal studies from the past, it would be wise to examine

development and child and family impacts over the entire early childhood period. For example, a study that combined Early Head Start and Head Start experiences over time would be fruitful. Looking at outcomes of children with disabilities throughout the early childhood period (e.g., combining Part C and Section 619 of IDEA) would be a more meaningful picture of child readiness for school.

Carole W. Brown,
Joneen M. Schuster, and Doyna Illmer

See also Child Abuse; Cognitive Development and School Readiness; Disabilities; Early Child Care and Education; Head Start; Individualized Education Program; Individuals with Disabilities Education Act; Measurement of Cognitive Development; Poverty; Risk Factors and Development; School Readiness; Special Education

Further Readings

Barnett, W. S., & Belfied, C. R. (2006). Early childhood development and social mobility. *The Future of Children: Opportunity in America*, *16*(2), 73–98.

Brooks-Gunn, J., Fuligni, A. S., & Berlin, L. J. (2003). *Early child development in the 21st century: Profiles of current research initiatives.* New York: Teachers College Press.

Gormley, W. T., Jr., Gayer, T., Phillips, D., & Dawson, B. (2005). The effects of universal pre-K on cognitive development. *Developmental Psychology*, *41*(6), 872–884.

Hebbeler, K., Spiker, D., Bailey, D., Scarborough, A., Mallik, S., Simeonsson, R., et al. (2007). *Final report of the National Early Intervention Longitudinal Study (NEILS).* Menlo Park, CA: SRI.

Minkovitz, C. S., Hughart, N., Strobino, D., Scharfstein, D., Grason, H., Hou, W., et al. (2003). A practice-based intervention to enhance quality of care in the first 3 years of life: The Healthy Steps for Young Children Program. *JAMA*, *290*(23), 3081–3091.

Office of Special Education Programs. (2006, July). *Twenty-sixth annual report to Congress on the implementation of the Individuals with Disabilities Education Act.* Retrieved March 28, 2007, from http://www.ed.gov/about/reports/annual/osep/2004/index.html

O'Neill, J., & Korenman, S. (2004, December). *Child poverty and welfare reform: Stay the course.* Retrieved April 1, 2007, from http://www.manhattan-institute.org/html/cr_44.htm

U.S. Department of Health and Human Services, Administration for Children and Families. (2005, June). *Head Start impact study: First year findings.* Retrieved March 21, 2007, from http://www.acf.hhs.gov/programs/opre/hs/impact_study/index.html

Wall, S. (2002). Opportunities and challenges in providing services to children with disabilities within Early Head Start. In Administration on Children, Youth and Families, *Making a difference in the lives of infants and toddlers: The impacts of Early Head Start* (Vol. 1: Final technical report). Washington, DC: U.S. Department of Health and Human Services.

Westrup, B. (2007). Newborn Individualized Developmental Care and Assessment Program (NIDCAP): Family-centered developmentally supportive care. *Early Human Development*, *83*(7), 443–449.

EATING DISORDERS

Eating disorders are psychological syndromes characterized by severe disturbances in eating behavior, typically in conjunction with intense fears of weight gain, distortion in one's perception of one's body shape and weight, and/or having a self-evaluation that is unduly influenced by body shape and weight. Eating disorders are of particular concern to educational psychologists because they typically onset in mid-to-late adolescence, and also because sociocultural and normative factors are strongly implicated in their etiology. In addition, some researchers believe that eating disorders may be best ameliorated through prevention programs targeted toward high-risk adolescents and delivered in educational settings. In this entry, the symptoms and diagnostic criteria for the two main eating disorder diagnoses—anorexia nervosa and bulimia nervosa—are presented first, followed by a description of the prevalence of, and gender differences in, these disorders and a description of a miscellaneous eating disorders diagnosis: eating disorder not otherwise specified. Next, the history, course, causes, prevention, and treatment of these disorders are reviewed. Finally, other eating disorder diagnoses are discussed, including tentative diagnoses that need further research, such as binge eating disorder, as well as feeding disorders of infancy and childhood.

Anorexia Nervosa

According to the *Diagnostic and Statistical Manual of Mental Disorders* (Fourth Edition, Text Revision) (*DSM-IV-TR*), the psychiatric diagnostic classification system most frequently used by clinicians and researchers in the United States, to receive a diagnosis of anorexia nervosa (AN), an individual must maintain a body weight that is significantly less than normal. The *DSM-IV-TR* specifically states than an individual with AN must weigh less than 85% of the

weight considered normal for the individual's age and height; however, the *DSM-IV-TR* does not refer clinicians to a specific source to establish normal weight and, in fact, encourages clinicians to consider normal weight in light of individual factors such as body build. The current *International Classification of Diseases* (ICD-10), an alternative diagnostic classification system, specifies a more stringent standard for abnormally low body weight in AN; specifically, a body mass index of 17.5 or less.

To receive a diagnosis of AN according to *DSM-IV-TR* criteria, in addition to low body weight, an individual must also display an intense fear of gaining weight or becoming fat and a disturbance of certain weight- and shape-related thoughts. For example, an individual with AN may misperceive her body shape or size (believing she is larger than she really is), may use her weight as the only or the most important criterion against which she evaluates her self-worth, or may deny the seriousness or dangerousness of her low weight.

For women old enough to menstruate, the *DSM-IV-TR* also requires amenorrhea (absence of menstruation over several menstrual cycles) to make a diagnosis of AN. The reason for including amenorrhea as a criterion is that women who are abnormally thin or who have an abnormally low percentage of body fat frequently stop menstruating. However, because many women take hormonal birth control medications that cause regular menstruation regardless of body composition, this criterion can be difficult for clinicians to assess. Research also suggests that some women, even at very low weights, continue to menstruate, and several studies have found that deleting amenorrhea from the criteria does not reduce the quality of the AN diagnosis. For this reason, many researchers believe that amenorrhea will no longer be a requirement for a diagnosis of AN in the future.

After deciding that a patient meets criteria for AN, clinicians using the *DSM-IV-TR* for an eating disorder diagnosis must assign the individual to a subtype of AN. There are two subtypes: a restricting subtype and a binge-eating/purging subtype. Individuals who fall into the binge-eating/purging subtype of AN regularly engage in either binge-eating behavior, purging behaviors, or both (described more fully in the next section). Individuals with the restricting subtype of AN do not engage in such behaviors. Some people believe that anyone who engages in binge-eating or purging behaviors must have bulimia nervosa (BN), but this is not the case. The main feature distinguishing someone

with AN from someone with BN is body weight, not behavior; individuals with AN are abnormally thin, whereas individuals with BN are normal weight or overweight.

Bulimia Nervosa

BN is characterized by recurrent episodes of binge eating, followed by inappropriate compensatory behaviors designed to prevent weight gain, such as self-induced vomiting, misuse of laxatives or other drugs, periods of fasting, or excessive exercising. In the *DSM-IV-TR*, binge eating is defined as eating in a discrete period of time (e.g., a 2-hour period) an amount of food that is definitely larger than what most people would eat under similar circumstances and feeling a lack of control over eating during the binge episode. Some researchers call this definition an "objective binge" because the amount of food being consumed is objectively large, such as eating a half-gallon of ice cream or an entire extra-large pizza at a sitting. However, people with BN are more likely than objective raters to rate *any* given eating episode as a "binge." Researchers have also observed that the type of food consumed (e.g., foods perceived to be high calorie or "sinful," such as potato chips) and the individual's mood while eating are more predictive of whether an eating episode will be described as a binge than the actual amount of food consumed. Thus, some researchers believe the subjective aspects of a binge— the feelings of loss of control and negative emotions— may be more important to a diagnosis of BN than the "objectively large" binge criterion. However, other researchers maintain that the presence of objectively large binges has diagnostic significance.

In addition to binging, to receive a diagnosis of BN, an individual must engage in recurrent and inappropriate compensatory behaviors designed to prevent weight gain. Many people are aware of self-induced vomiting and laxative abuse as examples of compensatory behaviors engaged in by individuals with BN, but individuals with BN may also misuse other drugs, such as diuretics, diet pills, or enemas, or they may engage in periods of fasting, or they may simply exercise excessively. Exercise is considered "excessive" when it interferes with other important activities, such as schoolwork or relationships; when it occurs at inappropriate times and places; or when the individual continues to exercise despite injury. Individuals must engage in both binges and compensatory behaviors,

on average, twice a week or more for 3 months to receive a diagnosis of BN. In addition, similarly to individuals with AN, individuals with BN must be unduly influenced by their shape and weight in evaluating themselves.

As with AN, BN is divided into two subtypes: a purging subtype, where the individual engages in self-induced vomiting or the misuse of drugs to prevent weight gain; and a nonpurging subtype, where the ways that the person has tried to prevent weight gain have included only other methods, such as fasting or exercise. Some people may believe that someone who does not use laxatives or self-induce vomiting cannot have BN, but this is not the case.

Gender Differences, Prevalence, and Subthreshold Cases

Both AN and BN affect women disproportionately; 90% of people who meet criteria for either diagnosis are female (hence the convention in this entry to refer to individuals as "she"). Men do acquire eating disorders, however, and men and women with eating disorders appear to have similar symptoms, similar risk factors prior to developing the eating disorder, and similar co-occurring disorders (such as depression). Compared to women with the disorder, men may have a later onset of their eating disorder symptoms and may be more likely to have engaged in a sport where weight control was emphasized or to have been overweight prior to the onset of the disorder. Some researchers believe that eating disorder rates in men may be increasing as sociocultural pressures on men to be thin and athletic-looking increase.

Even in women, AN is relatively rare, with a lifetime prevalence among women of 0.5%. Bulimia nervosa is somewhat more common, with a lifetime prevalence among women of 1% to 3%. However, these may be underestimates of the true prevalence of eating disorder pathology. One serious criticism of the *DSM-IV-TR* classification of eating disorders is that the majority of patients seen in clinics who present with eating disorder symptoms do not actually meet full criteria for either AN or BN. People who have serious eating disorder symptoms, but who do not meet full criteria for either AN or BN, are usually diagnosed with eating disorder not otherwise specified (EDNOS), a miscellaneous diagnostic category that includes people with a wide variety of eating disorders symptoms and symptom severity. Little is known about how people with EDNOS differ from people with AN or BN, but they can have symptoms equally as severe as those who meet criteria for the full disorder. In one study of hundreds of patients who had sought treatment at eating disorder specialty clinics in several states, the percentage of patients who were diagnosed with EDNOS instead of AN or BN ranged from 49% to 71%. In other words, the relatively low prevalence of AN and BN may not reflect the true number of people who are troubled by serious eating disorder symptoms.

History of Eating Disorders

As far back as the Middle Ages, writers have described occasional cases of individuals, often young women, deliberately engaging in self-starvation, typically as part of religious devotion. In 1689, the physician Richard Morton described two cases wherein a young person (he described both a male and a female patient) under significant psychological stress refused to eat and became abnormally thin. His medical interventions with these patients were apparently of little benefit, and his female patient died (it is unknown what happened to his male patient following treatment). However, these early cases, however much they may resemble AN, lack the preoccupation with weight, shape, and avoidance of fatness that is the hallmark of the modern disorder. A description of the disorder in its contemporary form, including the association between self-starvation and fears of fatness or disturbed body image, as well as psychological distress, began to appear in accounts of the disorder at the beginning of the 20th century. BN, in contrast, has a much shorter history. Although there exist a small handful of case reports of binge-eating and compensatory vomiting or laxative abuse as early as the 1940s, the disorder was not named until Gerald Russell described it as a variant of AN in 1979. At that point, there were sufficient cases to mount a research program on the disorder, but it seems clear that BN largely appeared as a new disorder in the latter half of the 20th century.

Course of Eating Disorders and Relations Between Disorders

Eating disorders are often medically serious and can become chronic, although in some cases they also

remit, with or without treatment. The onset of both AN and BN is typically in mid-to-late adolescence, although AN usually begins slightly earlier than BN, and an onset for either disorder in early adulthood is also not uncommon. With respect to AN, a total recovery, with normal body weight, normal body image, and no intense fears of weight gain, is rare. Fewer than 50% of individuals with AN achieve recovery, defined as normal or near-normal body weight, resumption of normal menstruation, and eating patterns free of severe dietary restriction, binging, or purging. Even among these "recovered" individuals, however, distorted body image, ongoing fears of weight gain, and other psychological symptoms of eating disorders often continue. Although treatment for AN can be lifesaving in the short term, long-term recovery rates are not markedly improved for treated versus untreated individuals. About one third of individuals with AN can be expected to improve somewhat, while still displaying serious eating disorder symptoms, whereas about 20% continue to meet full criteria for AN even years later. The mortality rate for individuals diagnosed with AN is extremely high: about 5% per decade of follow-up, higher than for any other mental disorder except heroin dependence. Most deaths associated with AN are either the result of suicide or medical abnormalities that arise as a consequence of starvation. Suicide risk for individuals with AN is higher than that for any other mental disorder.

The prognosis for an individual with BN is overall more encouraging, although the course of BN also is frequently chronic. Between one third and one half of individuals with BN continue to meet full or partial criteria for BN in long-term follow-up studies, although effective treatment (discussed in more detail later in this entry) appears to improve the course of BN. Less is known about the course of EDNOS; one study that followed individuals who met partial but not full criteria for eating disorders found that most still struggled with symptoms 3 years later. In fact, nearly half of these participants went on to develop either AN or BN by the end of the study. In other words, the EDNOS category includes individuals who do not necessarily have a less severe eating disorder and who may go on to develop either AN or BN.

As this discussion makes clear, moving between eating disorder diagnoses is common; people initially diagnosed with AN frequently go on to develop BN, although an individual with BN subsequently developing AN is somewhat less common. Also, many individuals with AN or BN improve somewhat while still retaining significant eating disorder symptoms, thus meeting criteria for a subthreshold eating disorder, often classified as EDNOS.

Causes of Eating Disorders

Genetic Influences

Twin studies estimate the degree to which genes influence the development of disorders by comparing monozygotic (or identical) twins with dizygotic (or fraternal) twins to see if both members of the twin pair have a history of the disorder. When both members of a twin pair have (or do not have) a disorder, this is called *concordance*. If monozygotic twins are more likely than dizygotic twins to be concordant for a disorder, researchers consider this to be evidence that the disorder has a genetic component, because both monozygotic and dizygotic twins presumably share a family environment, but only monozygotic twins also share all of their genes. Twin studies in eating disorders find that concordance rates for monozygotic twins are higher than rates for dizygotic twins in AN, suggesting that vulnerability to AN is heritable. However, monozygotic and dizygotic twins have similar concordance rates in studies of BN, suggesting that a specific vulnerability to BN is probably not heritable.

Family studies, where researchers look for evidence that a given problem runs in families by researching the relatives of people who have a given disorder, have revealed that eating disorders run in families, such that families with one person who has AN are more likely to have other members who also have AN, BN, or other eating disorder pathology. In addition, in the case of individuals with either AN or the binge-purge subtype of BN, family members are more likely to suffer from substance abuse or dependence, as well as depression. Viewed as a whole, the family and twin studies of eating disorders suggest that there may be a specific genetic vulnerability for AN and a more general genetic vulnerability to psychopathology that influences a range of psychological problems, including BN.

Family Influences

Certain specific aspects of a child's upbringing or life experiences prior to developing an eating disorder have been implicated as risk factors in eating disorders.

These include some eating and appearance-related factors, such as having parents who are themselves dieting or who are critical of the child's appearance. In addition, family risk factors less obviously related to appearance or food have also been identified, including parents who are frequently absent or underinvolved in the child's life, parents with high expectations for the child or who frequently criticize the child, and discord or conflict between the parents. Experiences of abuse have also been implicated as risk factors for eating disorders; sexual abuse has been most frequently researched as a risk factor, but physical abuse and neglect of a child also appear to be important.

Dieting

Interestingly, the practice of dieting may contribute to eating disorders, specifically by increasing the risk of symptoms of binging and purging. Early indications that dietary restriction could lead to eating disorder symptomatology came from a famous study conducted by physician Ancel Keys and his colleagues during World War II. The goal of the study was to understand how to best treat starvation in war victims, and the research was conducted on 40 men who had been pre-screened for exceptional physical and psychological health. During the study, these men underwent a 24-week period of semi-starvation. Many of the men developed severe psychological symptoms and abnormal eating patterns during the course of study, including binge-eating, food stealing and hoarding, severe depression, and even psychotic symptoms. A remarkable 50-year follow-up found that fully 25 of the original 40 participants (all who consented to be interviewed) reported ongoing abnormal eating behaviors and ruminations related to food that had persisted even a half-century later.

The role of dieting in abnormal eating behaviors is further supported by experimental studies, which have demonstrated that when dieters are given a high-calorie food, they tend to consume more of it than do nondieters. They also fail to regulate their food consumption in an experimental setting; whereas nondieters will eat less of a high-calorie food if they have eaten a high-calorie food earlier, dieters eat more of the high-calorie food than nondieters do on both occasions. Such research suggests that dietary restriction under some circumstances may encourage binge behavior, which then may encourage purging. Several longitudinal studies have provided support for this view by establishing that dieting behavior predicts later onset of eating disorders, particularly BN.

Sociocultural Factors

One major factor implicated in the etiology of eating disorders is a cultural emphasis on the "thin ideal," or an unattainably slender figure as a standard of beauty (especially female beauty, although many researchers believe that cultural pressures on men to meet an unattainable physical ideal are also increasing). Eating disorders are rare or unknown in cultures that do not place an emphasis on thinness as a virtue or a standard of beauty; research suggests that eating disorders increase in prevalence when cultures begin to adopt a thin ideal or among individuals migrating from a culture that does not endorse the thin ideal to one that does. Eating disorders also appear to have increased dramatically in prevalence in Western countries as thinness has increased as a cultural ideal. For example, examining the weight and body size of both Miss America winners and *Playboy* centerfolds reveals a trend for decreasing weight over the past 50 years in women featured in these venues. Weight loss articles published in magazines have also increased over the same time frame, as has the prevalence of eating disorders. Some researchers believe that BN, not described until 1979, is more causally linked to sociocultural pressures than is AN, which has appeared in some form in many time periods and cultures and which appears to be more strongly heritable. Experimental research varying participants' exposure to the thin ideal (by having some participants look at pictures of thin models or receive fashion magazine subscriptions) also suggests that participants' body dissatisfaction and internalization of the thin ideal (as well as feelings of depression) increase with exposure to thin ideal media images. Longitudinal research shows that body dissatisfaction and thin-ideal internalization, in turn, are related to increased eating disorder symptoms.

Many researchers, however, have observed that whatever the overall influence of sociocultural pressures on eating disorder prevalence and symptoms, exposure to social pressures to be thin is not sufficient to cause eating disorders. If it was, most people exposed to such pressures—most people living in Western cultures—would have eating disorders. This is not the case. Researchers therefore have developed several integrative causal models of eating disorders that typically combine biological influences, individual

characteristics, familial and social pressures, and dieting behaviors to predict the acquisition of an eating disorder. Researcher Eric Stice and his colleagues, for example, developed an influential dual-pathway model of BN that posits that bulimic symptoms are caused by dieting, negative emotions, and the interaction between them, which in turn are caused by body dissatisfaction caused by pressures to be thin and the degree to which one has internalized the thin ideal. Other influential models emphasize biological and genetic characteristics, personality characteristics (such as perfectionism or need for control), or cognitions related to weight and shape.

Prevention and Treatment

Eating disorders are difficult to treat. Although inpatient treatment of AN does result in temporary weight gain as a result of refeeding of patients while they are hospitalized, no treatment, medical or psychological, has definitive proof of efficacy in treating AN long term. A new form of family therapy for AN, the Maudsley approach, has shown a great deal of initial promise as an efficacious treatment for adolescents with AN, with as many as 90% of adolescents treated with this approach achieving long-term recovery from the disorder in clinical trials. However, as the treatment focuses primarily on empowering parents to work with their children in stopping the disorder, it may not be suitable for adults with AN, and research on this approach is still ongoing.

In contrast to AN, several treatment approaches are known to be efficacious in addressing BN. Specifically, both cognitive-behavioral therapy (CBT) and interpersonal psychotherapy (IPT) have good empirical support as efficacious treatments for the disorder. In CBT, clinicians focus on altering eating disorder behaviors and cognitions through challenging irrational beliefs related to shape and weight and through monitoring and changing binging and purging behaviors using behavioral techniques. In IPT, clinicians focus not on the eating disorder behaviors per se, but instead on helping patients identify and modify their current interpersonal problems. Both CBT and IPT appear to be equally efficacious in treating BN, although patients receiving CBT appear to improve more quickly, and CBT has been more extensively researched. Although both therapies are superior to other treatments for BN (such as antidepressant medication) for the average patient, only about half of the

patients treated with these treatments cease binging and purging. Thus, there remains a great need for further research to identify treatments that will help those who do not improve.

Because eating disorders can be so serious, and the currently available treatments are so limited, many researchers have focused attention on preventing instead of treating them. Prevention programs vary in their effectiveness. Ineffective prevention programs are primarily those which are shorter in duration (e.g., only one session in length), which present material in a didactic (rather than interactive) style, or which focus on educating participants about the disorders. More effective programs, in contrast, target individuals who are at high risk for the disorders, last longer than one session, and use interactive teaching methods. In addition, effective programs focus on helping participants change their level of body dissatisfaction or internalization of the thin ideal or other factors thought to cause the disorders, rather than merely providing participants with information about the disorders. Some of the most effective prevention programs actually do not describe themselves as eating disorder prevention programs at all but instead are described as body acceptance programs, emphasizing the focus on underlying causes of eating disorders. One promising and effective approach to prevention focuses on increasing participants' discomfort with the thin ideal by asking them to dissuade others from adopting this ideal. Although prevention approaches must be carefully designed to be effective, they show great promise in decreasing eating disorders. Many researchers believe that persuading media outlets to adopt a more realistic ideal of beauty might be the most effective means of prevention of all.

Other Eating Disorders

Although AN, BN, and EDNOS represent prototypical eating disorders, other problems related to eating have been identified. These include both *DSM-IV-TR* disorders, such as the eating disorders of infancy and early childhood, and disorders that have been designated as deserving of further study, such as binge eating disorder (BED). BED consists of recurrent and frequent episodes of binge eating as in BN. In BED, these episodes of binging are related to feelings of distress and other problems in living, such as obesity, but unlike in BN, the individual does not engage in any compensatory behaviors to prevent weight gain and may not

experience any distortions of body image or fears of fatness that are central to the other eating disorders. Little is known about BED compared with the other eating disorders, but there are important differences between BED and the other eating disorders. Individuals presenting for treatment of BED tend to be older than those with AN or BN (in middle age rather than adolescence) and are more frequently male (as many as 25% are male). In addition, although BED seems to be associated both with obesity and feelings of depression, the distress and risk associated with the disorder may be less acute than in AN or BN. In addition, it is not clear how individuals with BED who seek treatment differ from individuals who routinely overeat but do not experience distress related to overeating.

Even more distinct and unrelated to the prototypical eating disorders are the eating and feeding disorders of early childhood, which include pica, rumination, and feeding disorder of infancy or early childhood. Pica consists of repeatedly ingesting inedible substances, such as paint, plaster, sand, or soil, in a fashion that is inappropriate to the person's developmental level and culture. Rumination consists of repeatedly regurgitating and rechewing food after feedings. Feeding disorder of infancy or early childhood consists of a child younger than age 6 failing to eat adequately and consequently losing weight or failing to gain weight. None of these disorders is associated with fears of fatness or concerns about body image. In most cases, these disorders are associated with mental retardation or developmental delays, a lack of an appropriately enriching or nurturing environment, or child abuse or neglect.

Margit I. Berman

See also *Diagnostic and Statistical Manual of Mental Disorders*; Malnutrition and Development; Obesity

Further Readings

Dingemans, A. E., Bruna, M. J., & van Furth, E. F. (2002). Binge eating disorder: A review. *International Journal of Obesity*, *2*, 299–307.

Harris, E. C., & Barraclough, B. (1997). Suicide as an outcome for mental disorders: A meta-analysis. *British Journal of Psychiatry*, *170*, 205–228.

Herzog, D. B., Hopkins, J. D., & Burns, C. D. (1993). A follow-up study of 33 subdiagnostic eating disordered women. *International Journal of Eating Disorders*, *14*, 261–267.

Keys, A., Brozek, J., Henschel, A., Mickelsen, O., & Taylor, H. L. (1950). *The biology of human starvation.* Minneapolis: University of Minnesota Press.

Mitchell, J. E., Cook-Myers, T., & Wonderlich, S. A. (2005). Diagnostic criteria for anorexia nervosa: Looking ahead to DSM-V. *International Journal of Eating Disorders*, *37*, 95–97.

Polivy, J., & Herman, C. P. (2002). Causes of eating disorders. *Annual Review of Psychology*, *53*, 187–213.

Silverman J. A. (1997). Anorexia nervosa: Historical perspectives on treatment. In D. M. Garner & P. E. Garfinkel (Eds.), *Handbook of treatment for eating disorders* (2nd ed., pp. 3–10). New York: Guilford Press.

Stice, E. (2001). A prospective test of the dual-pathway model of bulimic pathology: Mediating effects of dieting and negative affect. *Journal of Abnormal Psychology*, *110*, 124–135.

Stice, E., & Shaw, H. (2004). Eating disorder prevention programs: A meta-analytic review. *Psychological Bulletin*, *130*, 206–227.

Wilson, G. T. (1997). Cognitive behavioral treatment of bulimia nervosa. *The Clinical Psychologist*, *50*, 10–12.

Woodside, D. B., Garfinkel, P. E., Lin, E., Goering, P., Kaplan, A. S., Goldbloom, D. S., et al. (2001). Comparisons of men with full or partial eating disorders, men without eating disorders, and women with eating disorders in the community. *American Journal of Psychiatry*, *158*, 570–574.

EDUCATIONAL TECHNOLOGY

Educational technology can be thought of as the hardware, software, and "thinkware" of learning, as well as related research theory that guides these efforts. This entry first defines what educational technology is. Next, it presents a short history of the field. Then it presents highlights of the hardware, software, and audiovisual technologies used by educational technologists; processes and theories used by educational technologists; and research in educational technology.

What Is Educational Technology?

Educational technology is "the study and ethical practice of facilitating learning and improving performance by creating, using, and managing appropriate technological processes and resources," according to the Association for Educational Communications and Technology (AECT), the largest professional society focused solely on educational technology.

In practice, that encompasses several distinct areas of effort. One is planning and installation of technology for education. Since the mid-1970s, this technology has increasingly meant the computers used to teach and manage learning and related activities, but it also includes other hardware and operating technologies, from film and audio equipment to LCD projectors and other sophisticated projection systems.

A second area of focus for educational technologists is educational software. Some educational technologists develop learning programs that can be taken on a computer. These programs can be *synchronous* (occur when learner and instructor are both online at the same time, such as in a chat) or *asynchronous* (instructor or learner are not online at the same time, such as in a self-study tutorial). Other educational technologists develop software tools that are used to develop these learning programs, such as *virtual classroom software* (like WebEx and Centra), used to provide synchronous learning experiences, and software for creating tutorials, called *authoring tools* (such as Dreamweaver Course Builder and Authorware). This software focus is reflected in the definition of educational technology offered by the International Technology Education Association (ISTE): "Using multimedia technologies or audiovisual aids as a tool to enhance the teaching and learning process."

A third focus of educational technologists is the processes and products used to design instruction. In these applications, technology refers to the sequence of activities and the applications of cognitive psychology to the design of instruction, rather than a physical hardware or a type of software. These activities are known as *instructional-design*, as phrased by Charles M. Reigeluth.

In terms of academic study, educational technology is available as an academic degree from universities in nearly every U.S. state, as well as in Australia, Canada, the United Kingdom, and several countries in the Middle East and Asia. In its most academic form, educational technology explores the following: the models and processes used to analyze, design, develop, deliver, implement, and evaluate instruction; the technology used to support these processes—both analogue and computer-based—in order to deliver learning materials, facilitate communication, and provide assessment and feedback; the policies associated with these practices; and the research that informs practice.

The academic study of educational technology explores learning throughout the life cycle, from preschool through retirement education. Similarly, because educational technology encompasses technology, learning, and design, it is an interdisciplinary field that draws upon other fields such as educational psychology, developmental psychology, early childhood and elementary education, higher education, adult education, computer science, ergonomics, sociology, and communications studies.

History

Educational technology emerged as a field in its own right in the 1920s, though some point to the initial development of human communications as one of the first developments in educational technology, because these facilitated the exchange of knowledge between members of tribes and from generation to generation. The invention of writing and writing instruments, such as papyrus and pens, further facilitated that process.

Modern educational technology charts its roots to the advances in media occurring in the late 19th century, such as inventions of visual and audio recording and playback devices. Soon afterward, each found its way into education, as educational films (the first appeared in the early 20th century), educational records, and the use of photographs in textbooks.

During the 1920s and through the 1940s, with advances in other media, such as audio recordings, radio, and motion pictures, the focus shifted from visual instruction to audiovisual instruction. This trend continued through the 1950s with the advancement of television. At this point in time, educational technology focused primarily on the use of instructional media, and the field was commonly known as *instructional technology*. In the 1950s, and most notably in the 1960s and 1970s, educational technology became more and more complex. It grew from a discipline concerned primarily with instruction using audiovisual aids to one concerned with systematic instructional design. At the same time, the field of education began to see instructional technology as a process that looks at instructional problems and explores solutions to these problems.

As technological advances exploded with the computer revolution, educational technology became a complex, integrated process that became increasingly professionalized. The AECT developed a new definition of the field in 1977 that included the importance of the planning and analysis phases of instructional design. This definition also included a list of learning resources and clarified terms such as learning problems

and solutions. This new definition emphasized that instructional technologists were interested in more than just the development and use of media. The discipline began to attract individuals from various disciplines.

With the growth of the microcomputer, Internet, interactive video, and other communication technologies, educational technology has become accessible to the general public and to most organizations. In addition, behavioral learning theories, which previously served as the basis of instructional design, gave way to cognitive and constructivist theories. These developments led to a growing interest in the development of instructional strategies that incorporated these new developments. As a result, in the mid-1990s, a new definition of educational technology emphasized the main areas of study and practice in the field; design, development, utilization, management, and evaluation. The interrelationship of these domains emphasizes the link between teacher, learner, and instructional method. The profession is now more concerned with the application of a variety of instructional techniques, ranging from traditional lecture approaches to open-ended distance learning environments.

In the new millennium, the field has continued to grow and change as society and learners evolve. The number of universities and the expansion of mass education along with the inadequate funding provided by governments, as well as the increasingly heterogeneous background of learners, are just some of the challenges. Increasing numbers of knowledge workers are seeking jobs that correspond to new job requirements. With the Internet, there has been an explosion in the volume of information and potential knowledge available. All of these advances have led to a more individualized type of education that is geared to specific needs, an education that is sometimes distance-oriented but primarily subject-specific. As technology becomes increasingly complex, there is a greater need for professionals who are able to train and help users become adept and skillful in the use of the new technologies.

Although many in the field continue to work on media and tasks related to the instructional process, many work in processes that are influencing the world as a whole. As the world outside the university changes dramatically, there is an increased demand for highly skilled workers, who need to be continually learning and improving their professional technical skills. As a result, educational technologists have been paying more attention to improving workplace performance and knowledge management. As the academic world gradually adapts to the real world of technological and societal changes and lifelong learning becomes critical to economic growth and development, the field of educational technology responds. There is an increased emphasis on independent learning, on developing social skills and teamwork, on adapting to change, and on developing thinking and problem-solving skills. As the volume of available information expands, it has become increasingly important to develop systems that can help people get at and process this information. Educational technologists are active in all these various areas.

Hardware, Software, and Audiovisual Technologies

One area of focus within educational technology is the applications of technology in learning. Technologies generally fall into these categories:

- Technologies for teaching and learning
- Technologies for facilitating communication among participants in the learning process
- Technologies for facilitating evaluation
- Technologies for managing learning activities

Because standardization of technology is essential to the smooth integration of technology in education, any discussion of technology also addresses standards.

Technologies for Creating Learning Materials

Some argue that the first educational technologies include writing systems, papyrus, and the writing instruments on which to place characters on the papyrus. One of the purposes of ancient documents was to record history and key lessons so that future generations might learn from them.

But the modern practice of educational technology begins with the introduction of audiovisual technology in the late 19th and early 20th centuries. The technologies can be categorized as follows:

Technologies for designing and developing the materials. Most of these technologies were developed for assisting with the creation of content. One of the earliest technologies was the typewriter, which enabled people to create content faster, more legibly, and more consistently. Later versions of typewriters included the ability to correct errors, change the look of the typeface, and save content, so revisions would not require retyping. The advent of computers brought

about word processing, software that lets authors write, change, revise, and print text; desktop publishing, which allowed the low-cost production of high-quality educational materials (before that, material was prepared by hand with highly skilled labor); and software for publishing materials online, including the HTML standard that lets Internet browsers like Mozilla Firefox and Internet Explorer display content produced from a variety of sources. Whereas none of these technologies was specifically produced for educational purposes, there was one that was: the authoring system, which is used to create online learning programs for people with little or no knowledge of computer programming. This software lets authors not only create the content but also determine what it looks like and test learners' understanding of the content. One of the first major authoring systems was PLATO; it appeared in the late 1960s and was intended for mainframe computers. In the mid-1980s, several key authoring systems for microcomputers emerged, including the Hypercard for the Apple computer, which also had other purposes and was intended for people with limited computer literacy, and Authorware, which was more sophisticated.

Technologies for producing the materials. This category includes equipment for recording visual images (photography, film, and video cameras, as well as related lighting) and audio (sound recorders, microphones). It also includes equipment for developing photographs and editing visuals and sounds together into a single, coherent product. In the past, this equipment worked with analog technology. Since the early 1980s, recording and production have converted to digital technologies. Note that these technologies were not developed specifically for use by educational technologists. However, the price consciousness of many educational technology departments provided incentives to manufacturers to develop more modestly priced versions of this equipment.

Technologies for reproducing the learning materials, which involves making copies of audiotapes, films, books, articles, and similar materials. These technologies were not developed specifically for educational technologists, and many educational technologists use outside companies to handle these tasks (and, therefore, have limited exposure to them).

Technologies for playing the recordings or using the content, which involves the use of specialized

projectors and similar playback equipment that often cost thousands of dollars. For audio, playback equipment includes record, tape, CD, and MP3 players. For visual recordings, playback equipment includes filmstrip and slide projectors. For combinations of audio and visual recordings, playback equipment includes film projectors and screens and video players and monitors. Audio and visual recordings have also been synchronized, such as slide-sound shows in which a slide projector automatically advances a slide show when it hears a special sound from a tape recorder. More recently, playback is available on computers, personal DVD players, and personal digital assistants, like PalmPilots and iPaqs. Playback equipment also includes equipment for projecting these visual images so large audiences can see them. Projection equipment includes screens, LCD projectors, large screen monitors, and rear-projection systems. Although none of these technologies was developed specifically for the educational environment, versions of the products were adapted specifically for educational environments.

Most of the technologies for creating learning materials require specialized knowledge. As a result, the curricula of many academic programs in educational technology programs include components on the use and operation of this equipment.

Technologies for Facilitating Communication Among Participants in the Learning Process

A second group of educational technologies is intended to facilitate communication between instructors and learners. These technologies can be categorized as follows:

Technologies for facilitating communication among the participants in the learning process. In the late 19th and early 20th century, these technologies primarily included mail and telephone. Both types of communication were intended for person-to-person communication. With the advent of computer-based technologies, several new options emerged. Some provided for person-to-person communication, including electronic mail (e-mail), Internet telephony, and instant messaging (IM). Other options offered by the computer included group-based communication. Early options included discussion lists (also called listservs, which are ongoing e-mail exchanges among a group of registered participants) and discussion boards,

programs that manage discussions on a class of topics. With the advent of social computing on the Web are wikis, which allow a group to rapidly create a work product together online, and blogs (short for "web logs"), which are like online columns from a newspaper or diaries, and to which readers can leave comments.

Technologies for replicating the classroom through technology. These technologies allow students to participate in class when they might not otherwise be able to, because of physical condition, time constraints, or geographic constraints. Early technologies for replicating the classroom include radio (such as the bush radio used in Australia) and conference calls by telephone. By the mid-20th century, television classrooms emerged (in fact, the use of television for learning was one of the reasons for establishing the U.S.-based Public Broadcasting System, as well as several academic educational technology programs, like the one at Concordia University in Montreal, Canada. Although some uses of television were based on public broadcast, private, closed circuit television systems also were used, such as the University of Minnesota's UNITE system and IBM's Interactive Satellite Education Network. Most closed circuit networks had two-way communication ability, which means that not only could students see and hear the instructor, but the instructor could see and hear the students. By the mid-1990s, software had been developed that allowed educators to create a classroom-like environment online. Called virtual classroom software, this software lets instructors speak, show images, and interact with students in large and small groups and lets students interact with their instructors.

Technologies for Facilitating Evaluation

A third group of educational technologies is intended to facilitate the process of evaluating the learning process. A technology from the latter half of the 20th century is Scantron, which lets instructors automatically grade exams with objective questions (true/false, multiple choice) by having students fill out a form that can be read by the grading machine.

Most of the developments in this area of educational technology occurred after computers came into common use. Some of these technologies focus on grading a particular assignment. One is software that conducts entire tests online. Original versions of this software limited educational technologists to objective questions, but later technologies have allowed instructors to

ask open questions (like essay questions) online. In most instances, instructors need to grade those answers themselves. Although some software has been developed to grade open questions, it is not yet in common use. A related technology lets instructors assess the likelihood that an essay or similar written assignment is plagiarized.

Other computer-based technologies for evaluating learning focus on broader issues. Portfolio software lets learners create a process portfolio, which allows them to include a class assignment and comment on what they learned from the assignment. An instructor, advisor, parent, or peer can review the assignment and reflections and offer feedback. Another is analytics software, which lets an administrator or manager compile statistics from a variety of computer-based sources, such as information about students, enrollment, and grades, and identify larger trends in the operation. Some of these analyses are presented as dashboards (e.g., the analysis of the American Society for Training and Development), which look like the dashboard of a car with a large number of indicators. When a problem seems to be arising in one of the indicators, the system visually signals the manager by turning that indicator red or orange.

Technologies for Managing Learning Activities

A fourth group of educational technologies is intended to manage learning activities. These technologies assist with the management of class activities as well as broader activities of the organizations that sponsor learning. Technologies for managing learning activities include learning management systems, course management systems, and learning content management systems.

Learning management systems. Learning management systems (LMSs), which act as electronic registrars, perform some or all of the following tasks for both online and classroom courses:

- Registration
- Tracking of participation (classroom attendance, sign-ons and sign-offs of online courses)
- Tracking of completions (including final scores or grades)
- Testing
- Providing aggregated reports, such as the numbers of people registered for particular courses

- Sharing information with other systems, such as human resource information systems
- Suggesting curricula for learners
- Tracking skills of individuals and groups of people within an organization
- Providing learners with a one-stop shop for their learning needs (especially online programs) and personalized information, such as a customized learning page that includes recommended paths through learning material

Examples of LMSs include Saba, Plateau, and NetDimensions EKP.

Course management systems. Course management systems (CMSs) were originally designed to support classroom learning in academic settings, such as universities and high schools. CMSs provide instructors with the ability to do the following:

- Place course materials online. Most CMSs provide preprogrammed buttons for the course syllabus, course schedule, and course materials linked to specific lessons, such as copies of readings and Power-Point slides from lectures.
- Track student progress, through assessment features (which let instructors give quizzes and tests online) and an online gradebook, where instructors can post student grades.
- Discuss readings with students and continue class discussions between formal class sessions, through the use of a discussion board.
- Send announcements to classes and communicate individually with students.
- Provide a lock box for students, where students can store class materials (e.g., a presentation to give later in class).
- Track course statistics, such as who used the course site and when.

Examples of CMSs include the commercial products Blackboard and WebCT and the open source system Moodle.

Learning content management systems. Learning content management systems (LCMSs) are software for creating, storing, retrieving, changing, and reusing material intended for use in an online learning program. LCMSs also store test questions and can be used to track activities in developing material. Finally, LCMSs can format the same content for presentation online and in print, and with a number of word processes. Examples of LCMSs include Eedo Force 10

and OutStart. Note that many LCMSs also have the registration and management functions of an LMS.

Standards for Learning Technology

Standards refer to rules that have been set up by various professional and trade organizations to address a variety of aspects of the development and production of learning content, so that content produced in one organization and on one brand of equipment can be easily used in another organization on another brand of equipment. Although all types of learning technologies involve the use of standards (e.g., how to take a photograph and ensure that the slide projector plays it), standards have only become a significant concern to educational technologists with the advent of online learning.

In technology for learning, a standard represents the last phase in a series of milestones intended to develop an idea for coordinating interactions among software and devices. The first phase results after someone or a group of people recognize a situation in which some coordination of efforts will result in more efficient or economical interactions among programs or some similar benefit to users. In many ways, this idea is similar to that proposed by Geary Rummler and Alan Brache for human performance technology. They note that efficiencies and improvements are often achieved by examining and reengineering the interactions among groups within and across organizations.

Processes and Theories

The field of educational technology is guided by a family of processes that form the core of most academic programs and are widely acknowledged by most practitioners. These processes are also part of introductory training to people who work in the field but do not have academic backgrounds in the field.

The core process followed by educational technologists is known as *instructional systems design*, which is the recommended process used to design, develop, and implement learning programs. Although more than 40 versions of this process exist, the core of each is the same and is known as ADDIE, for analysis, design, development, implementation, and evaluation,

Analysis. To identify the desired outcomes, the current performance, and the reason that the gap between current and desired performance exists. Analysis also identifies the intended learners, as well as the

conditions that affect the design, development, and delivery of the course, such as a tight deadline and a preference for learning in a particular format. The products of the analysis process are a set of behavioral objectives, which define the learning outcomes, and a test or similar assessment, which can be used to determine the extent to which learners have mastered those objectives.

Design. To turn the requirements into a learning program. During design, the educational technologist determines the medium of instruction—such as classroom or computer-based tutorial—and the format of the course. The educational technologist also determines the learning strategy to apply—that is, the approach to presenting and reinforcing the learning content, such as a mastery or discovery approach to learning. In many organizations, educational technologists also prepare storyboards (initial drafts) of learning materials at this time.

Development. To convert the design intentions into a working learning program. The actual activities vary, depending on the type of course, but usually include preparing several drafts of the content and obtaining reviews by subject-matter experts and editors and other educational technologists, and piloting tests (i.e., trial uses of the content) with intended learners.

Implementation. To make the learning program available to its intended learners. The actual activities vary, but they usually include announcement of the availability of the program, registration, and offering of the learning program. Implementation also encompasses the activities involved in operating a learning program, such as setting up a classroom or providing technical support for computers on which tutorials are taken.

Evaluation. To determine the extent to which the learning program achieved its objectives.

Although some refer to ADDIE as a model, it is neither descriptive nor predictive of actual practice. Several empirical studies suggest, instead, that although ADDIE is an ideal process, practicing professionals do not follow it. Rather, ADDIE is prescriptive of ideal practice in the field.

Guiding each phase of ADDIE are additional processes, theories, and methods. For example, guiding the development of objectives are approaches popularized by Bob Mager in the 1960s. There are a number of approaches that guide design. Some pertain to the use of—and choice of—media, best characterized by

the debate between Richard Clark and Robert Kozma on the role of media in instruction. Some approaches focus on appropriate strategies for designing instruction. Among the most popular (but certainly not an exhaustive list) are Robert Gagne's nine events of instruction; John Keller's attention, relevance, confidence, and satisfaction (ARCS) model; and problem-based learning (D. H. Jonassen). Guiding the development of learning programs are such approaches as advance organizers, theories of document and information design, and even the appropriate use of audio and visual cues. Different approaches guide evaluation, depending on the level of education. Guiding most of them is the belief that assessment should focus on the extent to which learners can accomplish the learning objectives (known as *criterion-referenced testing*). Although high-stakes testing—that is, standardized tests to assess the performance at a particular grade level—is the type of assessment best known in primary and secondary education, using portfolios to help both learners and teachers evaluate learning is gaining much interest among educational technologists. In terms of evaluating training, Donald Kirkpatrick's four-level approach is the most dominant. It evaluates satisfaction, learning, transfer of learning to on-the-job behavior, and the impact of the learning on the organization.

In addition, some educational technologists are guided in their work by the principles of human performance technology (HPT). HPT frames most issues addressed by its practitioners as performance problems. Such problems might result from a lack of knowledge and skills and can be solved by learning programs. But performance problems might also emerge from a lack of resources (e.g., receiving training on a new software program but not having access to that program) or motivation (e.g., fearing the loss of a job resulting from new technology). Learning programs will not solve problems resulting from a lack of resources or motivation; these problems require different types of interventions. HPT emerges from the workplace. Although its advocates claim that HPT applies to all of the environments that educational technologists address, the overwhelming body of evidence is focused on the workplace.

Research and Theory

Research in educational technology has four key streams. The first stream of research focuses on the *design* of instruction. Early research tended to focus on

the design of individual lessons and courses. Some studies focused on validating a particular design strategy. For example, studies in the 1950s and 1960s focused on applications of programmed instruction, a type of instruction that breaks learning down to its smallest tasks and involves teaching a topic one micro task at a time. Other studies compared instruction in different media, such as online versus classroom instruction. The dominant research methods used for both types of studies was experimental (either control group or quasi-experimental). In many cases, researchers present experience reports, that is, reports and some research results based on their own classroom experience. More recently, research in this area has focused on learning sciences, which, as defined by the International Society on Learning Sciences, uses "empirical investigation of learning as it exists in real-world settings and how learning may be facilitated both with and without technology." Studies explore issues of learning, teaching, and design and often employ qualitative research techniques.

The second stream of research is similar to the first but focuses on *applications of hardware and software technology* to learning. In the earliest days of the field, the research focused on the use of film and audio recordings for learning. As television came into wider use, research focused on its use. In fact, the use of television for learning gave a boost to the field, because many governments wanted to use television in teaching at all levels. Since the 1980s, most of the research has focused almost exclusively on uses of computer technology in learning. Some of the research attempts to provide proof-of-concept for specific technologies in well-defined learning contexts, such as uses of intelligent tutoring in mathematics and classroom response monitoring systems (in which students are provided with a key pad to type in responses to multiple-choice and similar questions) in large classes. Some of the research explores ways to use technologies, such as some of the recent research on mobile learning (learning using personal digital assistants, like PalmPilots) and podcasting (use of audio and video recordings broadcast on MP3 players like an iPod). Some of the research focuses on the technology used to manage learning, such as LMSs, which are widely used in the workplace to help workers enroll in and take courses, and CMSs, which are widely used in universities to help extend the classroom experience online by providing a place where instructors can distribute syllabi, assignments, and readings, as well as hold ongoing discussions with learners. This research often takes the form of experimental and quasi-experimental studies, but similar to research on the design of instruction, it increasingly incorporates qualitative techniques.

The third stream of research focuses on *resources* that complement the learning process. Often, these studies explore phenomena that are both a type of learning program and a type of software. For example, according to David Wiley, one area of great interest in the first half of the 2000 decade is reusable learning objects, which are materials designed for learning in one context but available for use in others. Also of great interest are portfolios, especially electronic ones, which, as mentioned earlier, are used for evaluation of learning.

The last stream of research focuses on the practice of *educational technology*. Most of this research focuses on the practice of instructional design. Some studies focus on which skills instructional designers use on the job; others focus on how extensively instructional designers practice ADDIE. Most of these studies are survey-based, but some of the more extensive ones use methodologies like developing a curriculum (DACUM) and are the basis of certifications in the field. In addition, some recent studies partially or completely employ qualitative techniques to compile rich description of the environmental factors affecting the design of learning.

Regardless of the stream of research, most research in educational technology tends to have a "micro" focus, that is, a focus on individual learning situations, such as a comparison of a course taught in two different environments or the application of a particular technology to a particular learning problem, such as the design of an intelligent tutoring system for solving a particular type of mathematical problem.

The early use of the experimental technique in research on educational technology is representative of the field's roots in behaviorism. Indeed, many of the early researchers in educational technology, such as Tom Gilbert and Robert Gagne, had worked with B. F. Skinner. But the field has increasingly embraced cognitivism and constructivism, as researchers become aware of the situational nature of learning and want to find stronger ways to reflect that in the research. One of the many responses to this shift is design research, which attempts to study the development of a course and use the research to inform its ongoing development.

Saul Carliner, Ofelia Ribiero, and Gary Boyd

See also Assessment; Curriculum Development; Learning

Further Readings

Association for Educational Communications and Technology (AECT) Task Force on Definition and Terminology. (1977). *The definition of educational technology.* Washington, DC: Author. (ERIC Document Reproduction Service No. ED192759)

Reiser, R. A., & Ely, D. P. (1997). The field of educational technology as reflected in its definitions. *Educational Technology, Research and Development*, *45*(3), 63–72.

Web Sites

Association for Educational Communications and Technology: http://www.aect.org

International Technology Education Association: http://www.iteaconnect.org

EFFECTIVE TEACHING, CHARACTERISTICS OF

Over the past several decades, a revolution has occurred in the definitions of "good" teaching. Researchers have found that defining good teachers by community ideals, personality characteristics, number of credit hours earned, or college grade point average proved disappointing, as these variables showed little relationship to what teachers actually do in the classroom. This directed researchers to study the impact of specific teacher activities on the specific cognitive and affective behaviors of their students. The term *good teaching* changed to *effective teaching*, and the research focus shifted from studying teachers exclusively to studying teachers *and* their effects on students. This new approach to studying classroom behavior has made the teacher–student relationship in the classroom the focus of modern definitions of effective teaching.

Linking Teacher Behavior With Student Performance

During this new era of research in classrooms, researchers developed methods of studying the classroom interaction patterns of teachers and students. Their goal was to discover which patterns of teacher behavior promoted desirable student outcomes. With a rich variety of classroom observation instruments, a picture of classroom activity could be captured within and across research studies and could be related to various measures of school achievement, such as better scores on classroom and standardized tests, an increase in the incidence of successful problem solving, and improved learning skills.

It was in this manner that consistent patterns of effective teaching began to emerge in studies conducted by different researchers. As in all research, some studies provided contradictory results or found no relationships among certain types of classroom interactions and student outcomes. But many studies found patterns of teacher–student interaction that consistently produced desirable student outcomes. The following are some of the most important characteristics of effective teaching to emerge from this research.

Key Behaviors Contributing to Effective Teaching

From this research approximately 10 teacher behaviors have shown promising relationships to desirable student performance, primarily as measured by classroom assessments and standardized tests. Five of these behaviors have been consistently supported by research studies over the past three decades. Another five have had some support from the research findings and appear logically related to effective teaching. The first five are key behaviors, because they are considered essential for effective teaching. The second five are helping behaviors that can be used in combinations to implement the key behaviors. The five key behaviors for effective teaching are (1) lesson clarity, (2) instructional variety, (3) teacher task, (4) student engagement in the learning process, and (5) student success rate.

Lesson Clarity

Lesson clarity refers to how clear a teacher's presentation is to the class, as indicated in the following examples:

More Effective Teachers

- Make their points clear to learners who may be at different levels of understanding
- Explain concepts in ways that help students follow along in a logical step-by-step order
- Have an oral delivery that is direct, audible to all students, and free of distracting mannerisms

Less Effective Teachers

- Use vague, ambiguous, or indefinite language
- Use overly complicated sentences that convey more than a single thought at a time

- Give directions that often result in student requests for clarification

One of the findings from research on lesson clarity is that teachers vary considerably on this behavior. Not all teachers communicate clearly and directly to their students without wandering, speaking above students' levels of comprehension, or using speech patterns that impair their presentation's clarity. A teacher with a high degree of clarity spends less time going over material and has more questions answered correctly the first time through the content, allowing more time for instruction.

Clarity is a complex behavior because it is related to many others, such as a teacher's organization of the content, lesson familiarity, and delivery strategies (e.g., whether a discussion, recitation, question-and-answer, or small-group format is used). Nevertheless, research shows that both the cognitive clarity and oral clarity of presentations vary substantially among teachers. This, in turn, produces differences in student achievement.

Instructional Variety

Instructional variety refers to a teacher's variability or flexibility of delivery during the presentation of a lesson. One of the most effective ways of creating variety during instruction is to ask questions. Many different types of questions can be integrated into the pacing and sequencing of a lesson to create meaningful variation. Therefore, the effective teacher needs to know the art of asking questions and how to discriminate among different question formats—fact questions, process questions, convergent questions, and divergent questions.

Another aspect of variety in teaching is perhaps the most obvious: the use of learning materials, equipment, displays, and space in the classroom. The physical texture and visual variety of the classroom contributes to instructional variety. This, in turn, promotes student engagement in the learning process and student achievement on end-of-unit tests. For example, some studies found the amount of disruptive behavior to be less in classrooms that had more varied activities and materials. Others have shown variety to be related to student attention.

Teacher Task Orientation

Teacher task orientation is a key behavior that refers to how much classroom time the teacher devotes to the task of teaching an academic subject. The more time allocated to the task of teaching a specific topic, the greater the opportunity students have to learn.

Some task-related decisions that a teacher must make are

1. How much time do I spend planning for teaching and getting my students ready to learn?

2. How much time do I spend presenting, asking questions, and encouraging students to inquire or think independently?

3. How much time do I spend assessing my learners' performance?

These questions pertain to how much material is presented, learned, and assessed, as opposed to how much time is delegated to procedural matters (e.g., taking attendance, distributing handouts, collecting homework, checking for materials). All teachers need to prepare their students to learn and want them to enjoy learning. However, most research studies agree that student performance is higher in classrooms with teachers who spend the maximum amount of time available teaching subject-specific content as opposed to devoting large amounts of time to the process and materials needed to acquire that content. It follows that classrooms in which teacher–student interactions focus more on subject-matter content that allows their students the maximum opportunity to learn and to practice what was taught are more likely to have higher rates of achievement. But, these classrooms also are those in which a dynamic relationship between teacher and learner provides the energy that motivates and challenges the learner to reach increasingly higher levels of understanding.

Student Engagement in the Learning Process

Student engagement in the learning process—or *engaged learning time*—is a key behavior that refers to the amount of time students devote to learning in a classroom. Student engagement is related to, but different from, a teacher's task orientation. Distinct from a teacher's task orientation—or the amount of time devoted to teaching a topic—is the time students are actively engaged in learning the material being taught. This has been called their *engagement rate*, or the percentage of time devoted to learning when students are actually on task, engaged with the instructional

materials, and benefiting from the activities being presented. Even though a teacher may be task oriented and providing maximum content coverage, some students may be disengaged. This means they are not actively thinking about, working with, or using what is being presented.

Such disengagement can involve an emotional or mental detachment from the lesson that may or may not be obvious. When students jump out of their seats, talk, read a magazine, or leave for the restroom, they obviously are not engaged in instruction. Students also can be disengaged in far more subtle ways, such as looking attentive while their thoughts are many miles away. An unpleasant fact of life is that a portion of every class may be off task at any one time, distracted for personal reasons and often amplified by an impending lunch period, a Friday afternoon, or day before a holiday. Correcting this type of disengagement may be much more difficult, requiring changes in the structure of the task itself and the cognitive demands placed on the learner. Researchers have contributed useful suggestions for increasing learning time and, more importantly, student engagement during learning. Their work has provided the following suggestions for teachers to promote student engagement:

1. Set rules that let pupils attend to their personal needs and work routines without obtaining permission each time.

2. Move around the room to monitor pupils' seatwork and to communicate an awareness of student progress.

3. Ensure that independent assignments are interesting, worthwhile, and easy enough to be completed by each pupil without teacher direction.

4. Minimize time-consuming activities, such as giving directions and organizing the class for instruction, by writing the daily schedule on the board. This will ensure that pupils know where to go and what to do.

5. Make abundant use of resources and activities that are at, or slightly above, a student's current level of understanding.

6. Avoid timing errors. Act promptly to prevent misbehaviors from occurring or increasing in severity so they do not influence others in the class.

Student Success Rate

A fifth key effective teaching behavior is student success rate. *Student success rate* refers to the rate at which students understand and correctly complete exercises and assignments.

A crucial aspect of the previously cited research on task orientation and student engagement is the level of difficulty of the material that was presented. In these studies, level of difficulty was measured by the rate at which students understood and correctly answered questions on tests, exercises, and assignments. Three levels of difficulty are

1. *High success*—the student understands the subject matter taught and makes only occasional careless errors;

2. *Moderate success*—the student has partial understanding but makes some substantive errors; and

3. *Low success*—the student has little or no understanding of the subject matter.

Not surprisingly, researchers have found that student engagement, that is, the time the learner is actively engaged with, thinking about, and working with the content being taught, is closely related to student success rate. Instruction that produces a moderate-to-high success rate results in increased performance, because more content is covered at the learner's current level of understanding.

The average student in a typical classroom spends about half of the time working on tasks that provide the opportunity for high success. But researchers have found that students who spend more than the average time in high-success activities have higher achievement, better retention, and more positive attitudes toward school. These findings have led to the suggestion that students should spend about 60% to 80% of their time on tasks that allow almost complete understanding of the material being taught with only occasional errors.

Summary of Five Key Behaviors

All five key behaviors—lesson clarity, instructional variety, teacher task orientation, student engagement, and success rate—are essential for effective teaching. Classroom researchers are studying other effective teaching behaviors and attaining a more thorough understanding of those already described. In other words, there can be no simple answer to the question, "What is an effective teacher?" Many activities must be orchestrated into patterns of behavior for teaching to be effective. The identification of only five

behaviors makes teaching appear deceptively simple. However, as the following section reveals, success in implementing these five key behaviors in the classroom will be assisted by other, helping behaviors. However, for the first time, research has provided a foundation for better definitions of effective teaching and for the training of teachers. Whereas these five behaviors constitute the skeleton of the effective teacher, there remains the heart, mind, and body to complete this picture of an effective teacher.

Helping Behaviors Related to Effective Teaching

To fill out the picture of an effective teacher, there are other behaviors that aid the teacher in performing the five key behaviors. These behaviors can be thought of as helping behaviors for performing the five key behaviors.

Research findings for helping behaviors, although promising, are not as strong and consistent as those that identified the five key behaviors. The research has not identified explicitly how these behaviors should be used. This is why helping behaviors need to be employed in the context of other behaviors to be effective, making them catalysts working in harmony with the five key behaviors rather than agents unto themselves. Among these helping behaviors are (a) using student ideas and contributions, (b) structuring, (c) questioning, (d) probing, and (e) teacher affect and the teacher–learner relationship.

Using Student Ideas and Contributions

Using student ideas and contributions includes acknowledging, modifying, applying, comparing, and summarizing student responses to promote the goals of a lesson and to encourage student participation. Note how any one of these activities could be used in achieving one or more of the five key behaviors:

- *Acknowledging:* Taking a student's correct response and repeating it to the class (to increase lesson clarity)
- *Modifying:* Using a student's idea by rephrasing it or conceptualizing it in the teacher's words or another student's words (to create instructional variety)
- *Applying:* Using a student's idea to teach an inference or take the next step in a logical analysis of a problem (to increase success rate)

- *Comparing:* Taking a student's idea and drawing a relationship between it and ideas expressed earlier by the student or another student (to encourage engagement in the learning process)
- *Summarizing:* Using what was said by a student or group of students as a recapitulation or review of concepts taught (to enhance task orientation)

More recently, the use of student ideas and contributions has been extended to reasoning, problem solving, and independent thinking. This has been achieved through teacher-mediated dialogue that helps learners construct or restructure what is being learned using their own ideas, experiences, and thought patterns. Teacher-mediated dialogue asks the learner not just to respond with a correct answer but also to internalize the meaning of what was learned by elaborating, extending, and commenting on it using the learner's own unique thoughts and experiences. In this manner, learners are encouraged to communicate the processes by which they are learning, thereby helping them to construct their own meanings and understandings of the content; this is called *constructivist teaching.*

Use of student ideas and contributions also can increase a student's engagement in the learning process. Thus, it has become a frequently used catalyst for helping achieve that key behavior. Research indicates that student ideas and contributions, especially when used in the context of the naturally occurring dialogue of the classroom, are more strongly and consistently related to student engagement than simply approving a student's answer with "good," "correct," or "right."

Structuring

Teacher comments made for the purpose of organizing what is to come, or summarizing what has gone before, are called structuring. Used before an instructional activity or question, structuring serves as instructional scaffolding that assists learners in bridging the gap between what they are capable of doing on their own and what they are capable of doing with help from the teacher, thereby aiding their understanding and use of the material to be taught. Used at the conclusion of an instructional activity or question, structuring reinforces learned content and places it in proper relation to other content already taught. Both forms of structuring are related to student achievement and are effective catalysts for performing the five key behaviors.

One method of structuring is to signal that a shift in direction or content is about to occur. A clear

signal alerts students to the impending change. Signals such as "Now that we have studied how the pipefish change their color and movements, we will learn . . ." help students switch gears and provide a perspective that makes new content more meaningful.

Another type of structuring uses emphasis. This method of structuring, called an *advance organizer*, helps the student to organize what is to follow. In this instance, students are asked to consider questions that might be raised and to know that generalizations beyond the concepts discussed will be expected. Phrases such as "Now this is important," "We will return to this point later," and "Remember this" are called *verbal markers* and can be used to emphasize important points.

In addition to verbal markers and advance organizers, the effective teacher organizes a lesson into an activity structure. An *activity structure* is a set of related tasks that increase in cognitive complexity and that, to some degree, may be placed under the control of the learner. Activity structures can be built in many ways (e.g., cooperatively, competitively, independently) to vary the demands they make on the learner and to give tempo and momentum to a lesson. For the effective teacher, they are an important means for engaging students in the learning process and moving them from simple recall of facts to higher response levels that require reasoning, critical thinking, and problem-solving behavior.

Questioning

Questioning is another important helping behavior. Few other topics have been researched as much as the teacher's use of questions. One of the most important outcomes of research on questioning has been the distinction between content questions and process questions.

Content questions. Teachers pose content questions to encourage the student to respond directly to the content taught. An example is when a teacher asks a question to see if students can recall and understand specific material. The correct answer is known well in advance by the teacher. It also has been conveyed directly in class, in the text, or both. Few, if any, interpretations or alternative meanings of the question are possible. Researchers have used various terms to describe content questions, such as direct, lower-order, convergent, closed, and fact questions.

Some estimates have suggested that up to 80% of the questions teachers ask refer directly to specific content and have readily discernible and unambiguous answers. Perhaps even more important is the fact that approximately the same percentage of teacher-made test items and behavioral objectives are written at the level of recall, knowledge, or fact. Therefore, test items, behavioral objectives, and most instruction seem to emphasize readily known facts as they are presented in curriculum guides, workbooks, and texts, leaving much less time for higher-order thinking, such as problem solving, decision making, and valuing. For the effective teacher, content questions are rarely ends in themselves but rather a means of engaging students in the learning process by getting them to act on, work through, or think about the material presented, leading to higher thought processes.

Process questions. To problem-solve, to guide, to arouse curiosity, to encourage creativity, to analyze, to synthesize, and to judge also are goals of instruction that should be reflected by questioning strategies. For these goals, content is not an end itself but a means of achieving higher-order goals. Researchers have used various terms to describe process questions, such as indirect, higher-order, divergent, open, and concept questions.

Process questions encourage thinking and problem solving by requiring the learner to use personal sources of knowledge to actively construct her or his own interpretations and meanings rather than acquiring understanding by giving back knowledge already organized in the form in which it was told. Process questions also help emphasize the learner's direct experience and the dialogue of the classroom as instructional tools while de-emphasizing lecturing and telling.

Probing

Another important helping behavior is probing, which refers to teacher statements that often follow questions that encourage students to elaborate on an answer, either their own or another student's. Probing may take the form of a general question or can include other expressions that elicit clarification of an answer, solicit additional information about a response, or redirect a student's response in a more fruitful direction. Probing often is used to shift a discussion to some higher thought level. Generally, student achievement is highest when the eliciting, soliciting, and (if necessary) redirecting occur in cycles. This systematically leads the discussion to a higher

level of complexity, as when interrelationships, generalizations, and problem solutions are being sought.

Teacher Affect and the Teacher–Learner Relationship

Anyone who has ever been in a classroom where the teacher's presentation was lifeless, static, and without vocal variety can appreciate the commonsense value of the affective side of teaching. However, unlike the behaviors discussed previously, affect cannot be easily captured in transcripts of teaching or by classroom interaction instruments. Consequently, narrowly focused research instruments often miss a teacher's affective nature, which emerges from a more holistic view of the classroom. This affective nature is the foundation from which effective teachers build a warm and nurturing relationship with their learners.

What the instruments miss, the students see clearly. Students are good perceivers of the emotions and intentions underlying a teacher's actions, and they often respond accordingly. A teacher who is excited about the subject being taught and shows it by facial expression, voice inflection, gesture, and movement, communicating respect and caring for the learner, is more likely to hold the attention of students and motivate them to higher levels of achievement than is a teacher who does not exhibit these behaviors. As a result, students will be more willing to actively engage themselves in the learning process.

Research has indicated that students take their cues from these affective signs and lower or heighten their engagement with the lesson accordingly. Enthusiasm is an important aspect of a teacher's affect in that it represents the teacher's vigor, power, involvement, excitement, and interest during a classroom presentation and his or her willingness to share this emotion with learners, who will want to respond in kind. Enthusiasm is contagious and can be displayed to students in many ways, the most common being vocal inflection, gesture, eye contact, and animation, but most importantly, how the teacher coordinates these signs to communicate that she or he cares about and respects the experiences, knowledge, and understandings students bring to the classroom. A teacher's enthusiasm has been found to be important in promoting student engagement in the learning process and achievement.

As research has turned from studying the individual characteristics of teachers to their interaction patterns with students, key and helping behaviors have been identified that are related to desirable outcomes in learners as measured by classroom and standardized tests of achievement, student projects, and performance assessments. Although not the only behaviors that define effective teaching, these key and helping behaviors provide a foundation for identifying, orchestrating, and teaching effective patterns of teacher–student interaction that result in desirable learner outcomes.

Gary Borich

See also Constructivism; Expert Teachers; PRAXIS™; Teaching Strategies

Further Readings

Berliner, D. (1984). The half-full glass: A review of research on teaching. In P. Hosford (Ed.), *Using what we know about teaching* (pp. 51–77). Alexandria, VA: Association for Supervision and Curriculum Development.

Berry, B. (2003). *What it means to be a "highly qualified teacher."* Chapel Hill, NC: Southeast Center for Teaching Quality.

Borich, G. (2007). *Effective teaching methods* (6th ed.). Upper Saddle River, NJ: Merrill/Prentice Hall.

Brophy, J. (2002). Teaching: Educational Practices series—1. United Nations Educational, Social, and Cultural Organization (UNESCO). Geneva, Switzerland: International Bureau of Education.

Darling-Hammond, L., Bransford, J., LePage, P., Hammerness, K., & Duffy, H. (Eds.). (2005). *Preparing teachers for a changing world: What teachers should learn and be able to do.* San Francisco: Jossey-Bass.

Fisher, C., Filby, N., Maliave, R., Cahen, L., Dishaw, M., More, J., et al. (1979). *Teaching behaviors, academic learning time and student achievement: Final report of phase III-B, beginning teacher evaluation study* (Tech. Rep. No. V-1). San Francisco: Far West Laboratory for Educational Research and Development.

McDermott, P., & Rothenberg, J. (2000, April). *The characteristics of effective teachers in high poverty schools: Triangulating our data.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA. (ERIC Document Reproduction Service No. ED442887)

Richardson, V. (Ed.). (2001). *Handbook of research on teaching* (4th ed.). Washington, DC: American Educational Research Association.

Rosenshine, B., & Stevens, R. (1986). Teaching functions. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 376–391). New York: Macmillan.

Wayne, A., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research*, *73*(1), 89–122.

EGOCENTRISM

A child playing hide-and-seek runs to the corner of the room and covers her eyes with a pillow while her father counts to 10, believing that if she cannot see her father, then her father cannot see her. A thirsty grocery shopper spends more money than usual on bottled water and decides not to buy the salty pretzels that are on his shopping list. Political conservatives responding to an Internet poll estimate that there are more like-minded conservatives in the population than political liberals. All of these people are reasoning egocentrically. That is, they are all reasoning about the world from their own current perspectives. In the first place, people process incoming information from their own perspective. Then, they extrapolate in two ways: temporally, extending present knowledge to the past and future; and socially, extending self knowledge to others. Education involves acquiring knowledge and communicating that knowledge to others, and egocentric bias in judgment is therefore a central—but easily overlooked—issue for educators.

History

Jean Piaget is credited with the classic demonstration of egocentric reasoning in children. In one version of his experiments, children viewed a three-dimensional mountain scene that included a house that was visible from their perspective but occluded from the imaginary perspective of a stuffed bear on the other side of the scene. The children indicated that the house could also be seen from the bear's perspective, leading Piaget to conclude that children of a certain age are unable to recognize that other beings can have differing points of view. Subsequent experiments revealed potential methodological flaws in the mountain task, but improved paradigms led researchers to draw essentially the same conclusions.

Since Piaget's demonstrations, research on egocentrism has advanced in various ways. First, egocentric biases have been widely demonstrated in adults. Like children, adults readily activate egocentric knowledge when rendering judgments, but unlike children, adults learn to correct egocentric knowledge to accommodate another person's perspective. Such corrective procedures require time, mental effort, and awareness of the need to correct. Insufficient correction is a common cause of egocentric biases in judgments among adults. Second, research on egocentric biases has expanded beyond visual perspective to incorporate perspectives based on knowledge, experience, culture, attitudes, and mood.

Consequences of Egocentrism

Egocentric reasoning can occur in at least three stages of processing and can, in turn, have consequences for both acquiring and communicating knowledge.

First, people must make sense of their environment by forming representations based on their perceptions. With visual information, for example, people are generally good at forming three-dimensional representations, based on their own singular perspectives, and inferring others' perspectives from this. Consequences may emerge, however, when they neglect to consider others' perspectives simultaneously. This can be illustrated by the presenter who repeatedly points at his own computer monitor, visible only to him, rather than gesturing toward the projection of this image, which is visible to the entire audience.

Second, people understand the past and the future according to their current perspectives. For example, people often use their current knowledge when judging the past. This leads to the logic referenced by aphorisms such as "hindsight is 20/20." When judging past decisions made under uncertainty, people will often rely on the outcome without appreciating that the outcome could not have been known at the time of the decision.

Finally, people think about others using their own perspective. This leads people to exaggerate the extent to which others share their beliefs, knowledge, and opinions. It also causes communication difficulties when speakers mistakenly assume that a listener shares one's knowledge when, in fact, they do not.

Importance for Education

Given the interpersonal nature of education, educators are at particular risk of many of the biases associated with the egocentric perception of other people. At least three potential difficulties can be identified, including motivating students, preparing lessons, and delivering information.

The tendency to exaggerate the extent to which others share one's beliefs and values, what psychologists call *the false consensus effect*, can make it difficult for educators to understand what will motivate their students. A teacher who imputes his own

preferences and values to students may be trying to motivate them with the wrong incentives.

An inherent difficulty for educators, who know the material they are trying to present, is to get beyond their own perspective to appreciate the perspective of students who do not know the material. Insufficient correction can cause teachers to overestimate how easy it is for students to learn the material and to overestimate the extent to which information has been clearly communicated—a problem psychologists call the "curse of knowledge." A math instructor writing a proof on the board might skip a number of steps because she assumes knowledge of algebra beyond the students' understanding. This can lead to confused students and a frustrated instructor.

Finally, effective communication requires instructors to be precise in language and clear in intentions to avoid ambiguities in the first place and to engage in perspective taking to resolve ambiguities that do arise. Psychologists have not yet uncovered a communication panacea, but a general awareness of potential pitfalls should be a valuable step toward more effective communication.

Given that people necessarily reason about the world from their own perspective, it is difficult to overcome egocentrism. A primary challenge for those in educational practice and research is to recognize the implications of this processing orientation, to diagnose it when it occurs, and to implement practices that help to overcome or circumvent negative consequences.

Benjamin A. Converse and Nicholas Epley

See also Piaget's Theory of Cognitive Development; Theory of Mind

Further Readings

Nickerson, R. S. (1999). How we know—and sometimes misjudge—what others know: Imputing one's own knowledge to others. *Psychological Bulletin*, *125*, 737–759.

Royzman, E. B., Cassidy, K. W., & Baron, J. (2003). I know, you know: Epistemic egocentrism in children and adults. *Review of General Psychology*, *7*, 38–65.

EMOTIONAL DEVELOPMENT

Emotional development refers to the process of learning to effectively express, regulate, and cope with one's emotions over time. Emotional development has significant implications for children's current and future functioning. Children who successfully navigate their emotional world are more likely to also be successful in their interpersonal relationships, academic and later employment endeavors, and in obtaining personal happiness and good adjustment. Children who encounter problems with their emotional development, on the other hand, are significantly more likely to have interpersonal difficulties, academic and later employment problems, and mental health or adjustment problems. This entry will first describe the important elements of emotional development: emotion reactivity, emotion regulation, emotionality, and emotional understanding. Next, an overview of the main influences on emotional development will be presented, including the child's own characteristics, interactions with parents, and exposure to broader family interactions. Finally, a discussion of how early emotional development is linked to future adjustment will follow.

Elements of Emotional Development

Emotional Reactivity and Expression

*Emotional reactivit*y refers to an individual's response to a stimulus change, or an alteration in the environment, which is reflected in physiological changes in the individual and is observed in the excitability or arousability of response systems, in terms of temporal features (e.g., how quickly the behaviors appear following the stimulus, how rapidly they escalate, how long they last, and how quickly or slowly they go away) and intensity features (e.g., how strongly the behaviors are expressed, how sensitive they are to stimuli). High emotional reactivity would be reflected in a child who begins crying and yelling vigorously at the first hint of a disagreement occurring between his or her parents, gets overaroused quickly, and remains so for an hour after the disagreement.

Early in infancy, basic emotions such as anger, sadness, and happiness begin to emerge. As children leave the toddler phase and enter preschool and grade school age periods, their emotional expression becomes more context dependent. For example, anger and happiness are expressed more than sadness and distress in peer settings. Furthermore, their emotional expressions become more complex over time, and they may begin to show blends of various emotions

and other secondary emotions, such as pride, shame, embarrassment, and guilt. These later-developing emotions reflect more self-consciousness and awareness and develop, not coincidentally, at the same time that children are beginning to have a sense of self and to be able to discriminate, compare, and appraise the self in relation to others.

Emotional Regulation

The second basic process involved in emotional development is *emotional regulation*, which is defined as the ability to modulate, control, or reduce the intensive and temporal features of an emotion. Regulation can occur at the neurophysiological, hormonal, attentional, and behavioral levels. Behaviorally, it is observed through approach or avoidance and inhibition responses. It is reflected in the ability to effectively maintain, enhance, or inhibit emotions appropriately. The term *coping* is sometimes used interchangeably with emotion regulation, as effective coping is inseparable from effective emotion regulation and vice versa.

As infants and toddlers, children require parental help in regulating their emotions, but as they grow older, they begin to develop their own strategies for regulation. Regulation strategies that children typically use, in order from most adaptive to least adaptive are problem solving (attempting to change the situation), support seeking from caregivers or peers (either for seeking solace or help), distancing-avoidance or distraction, self-calming behaviors (e.g., taking deep breaths), internalizing (e.g., keeping feelings to self), and externalizing or antisocial behaviors (e.g., hitting others or picking a fight). Using more adaptive strategies, such as seeking solace from others or self-calming strategies, is associated with more positive outcomes for children, such as greater self-worth. Although some children attempt externalizing or antisocial behaviors as means of emotion regulation, in fact those behaviors may actually reflect a lack of managing one's emotions. Although there is not a systematic empirical literature indicating what coping strategies emerge at what age, it is known that (a) as children age, they can generate more coping alternatives, and (b) older children are better able to utilize cognitively oriented coping strategies for situations in which they have no control.

Emotionality

The third basic process involved in emotional development is *emotionality*, which refers to an individual's predisposition toward more negative versus more positive emotions and reactions. Children with more negative emotionality are prone to more feelings of sadness, anxiety, and anger, whereas children with more positive emotionality may have a greater propensity for happiness and resilience.

Emotional reactivity, emotional regulation, and emotionality all interact together and the various combinations resulting from different levels of each of the three dimensions yield numerous possibilities for outcomes. Simply being high in negative emotionality may not necessarily result in negative outcomes for a child because the child may also be quite low in emotional reactivity (i.e., the child is more prone to negative feelings, but these negative feelings are not easily invoked), or the child may also be high in emotional reactivity but be quite effective in emotion regulation (i.e., the child is prone to more negative feelings and such emotions are quickly evoked and strongly expressed, but the child is able to just as quickly regulate his or her negative emotions and return to baseline). Alternatively, a child may be high in emotional reactivity and low in emotional regulation but have more tendencies toward positive emotionality, and thus, it is not problematic that he or she reacts quickly and vigorously and is not an efficient regulator because the highly positive emotions are not problematic or requiring regulation. Thus, emotional reactivity, emotional regulation, and emotionality may interact together to serve as either protective factors against emotional problems and disorders later in life or may serve as potential risk factors. Furthermore, emotional reactivity, emotional regulation, and emotionality are each thought to be influenced by both innate characteristics of the child and by social relationships, particularly parent–child relationships early in life.

Emotional Understanding

As children develop, they become more sophisticated in their *emotional understanding*. Young children not only show an awareness of their own emotional states, but they also become more adept at evaluating and appropriately responding to others' emotions. They also become increasingly better able to describe the causes and consequences of various emotions as they grow and become increasingly knowledgeable about emotional display rules (i.e., social customs for when, how, and to whom certain emotions are appropriate to express).

Children learn a great deal about emotional understanding through their own experiences of various emotions and the concurrent discussions that ensue about their emotional states. Parents are more likely to talk about feelings with their children during times when their children are experiencing distress or anger, and children, in turn, are more likely to engage in casual conversations about feelings when they are mildly upset. These casual conversations are important sources of learning for children and result in greater levels of emotional understanding later in life.

Influences on Emotional Development

Child Characteristics: Temperament

Temperament plays a fundamental role in current conceptualizations of emotional development and is one of the main contributors to children's emotional reactivity, emotional regulation, and emotionality. Temperament is typically defined as a set of inherited personality traits that are observable from the beginning of life. It refers to the child's organized style of behavior that appears early in development, such as fussiness or fearfulness, and shapes the child's approach to his or her environment and the environment, in turn, helps further shape the child's approach. Many different researchers have put forth different opinions regarding the various constructs that constitute temperament, including a broad range of individual characteristics, such as activity, sociability, impulsivity, mood, approach-withdrawal, intensity, threshold, rhythmicity, distractibility, attention span, persistence, and adaptability. Despite these differences, the dimensions that are consistent across definitions are those pertaining to emotions. Thus, temperamental differences in infants and children may be reflected in consistent individual differences in the amount and severity of expressed negative or positive affect, fearfulness, irritability, ease of being soothed, and withdrawal from, or attraction to, novel situations across a variety of contexts.

In their key longitudinal study on temperament, Stella Chess and Alexander Thomas describe four types of temperaments that children can be classified into: easy, difficult, slow-to-warm-up, and normal/moderate. The easy category represents about 40% of children and is characterized by children who are approachable, are adaptive to the environment, possess the ability to regulate basic functions (e.g., eating, sleeping, and elimination) relatively smoothly, display predominantly positive emotionality, and settle easily into new routines. The difficult category represents about 10% of children and is characterized by children who are predominantly negative or intense in mood, not very adaptable, and arrhythmic with regard to basic functions. Children with a difficult temperament may show distress when faced with new or challenging situations and/or may have a general distress proneness or irritability, especially when limitations are placed on them. These infants cry frequently in spite of their needs being met, seem to show no regular pattern for when they eat or sleep, and are difficult to soothe. The slow-to-warm-up category represents approximately 15% of children and is characterized by fearfulness or inhibition. These children are cautious in their approach to new or challenging situations. They are more variable in regulation and adaptability and may show distress or negativity toward some situations. They fit their namesake: They are slow to warm up to situations and people, but they do eventually adapt with effort. Finally, about 35% of children fall into the normal/moderate category, which is characterized by more typical babies, that is, ones who are more moderate and fall in between categories. They are not as easy as easy babies, but not as difficult as difficult babies. They may display some negativity, but not as consistently as those with difficult temperaments. They also may display some inhibition, similar to the babies in the slow-to-warm-up category, but again not as consistently. They also have more regularity to their biorhythms than those in the slow-to-warm-up or difficult categories.

Similar to these concepts of temperament, Jerome Kagan also described two temperament categories in children that represent extremes of behavior: inhibited versus uninhibited children. Inhibited children are characterized by shy, restrained behavior and fearfulness, whereas uninhibited children are characterized by sociable, bold, and approaching behaviors. Kagan views temperament traits as being based on qualitative and innate differences.

The key question of interest with regard to emotional development is what are the links between temperament and later risk for emotional problems or psychopathology? Numerous researchers have investigated this question and have found support for moderate links. In particular, having high emotionality as an infant has been linked to anxiety, depression, and attention problems in middle school. Falling in the

inhibited category as an infant is associated with displays of shyness and withdrawal in middle childhood, more social withdrawal in adolescence, and more dependence and introversion in adulthood. Falling in the uninhibited category as an infant, on the other hand, predicts more externalizing problems and increases in aggression later in childhood. However, temperament alone is not the sole risk factor for adjustment problems later on, and not all infants who have more problematic temperaments experience emotional difficulties later in life.

Regarding developmental outcomes and the risk for psychopathology, temperament itself is not necessarily the problem. The key issue is the goodness of fit between the parents and the child's temperament. Early in life, infants are not able to regulate their own emotions, and thus, it is the parents' task to help their infant regulate negative emotions and manage the intensity of positive emotions. Parents' responses to infants' temperament can help to attenuate or accentuate temperamental extremes. For example, if a child is highly reactive and intense and needs more care but the parent is able to effectively meet the child's needs and helps soothe and regulate the infant's negative emotionality, then the parent has successfully attenuated any potential problems arising from the more difficult temperamental behaviors. On the other hand, parents who do not successfully help their child in appropriate emotion expression and regulation, or even exacerbate it due to their own negative reactions to the child's behavior, may actually increase the risk of problems for their child. Thus, how parents interact with their child's temperament and the goodness of fit between the child's temperament and the parents' reactions are what matter most.

Relational Influences

The family is clearly the most important relational influence on children's emotionality and emotional development. Traditional research has typically focused on the importance of the parent–child relationship for children's emotional functioning. Although the parent-child relationship is the most significant single category of family influence on child development, other family relationships must also be considered to obtain a more comprehensive account of family influence on children's emotional development. Emerging research over the past two decades has highlighted the importance of also considering influences of the marital relationship

on children's emotional development. Although other factors (e.g., peers, schools, culture) also undoubtedly affect children's emotional development, a consideration of these additional influences is beyond the scope of this entry.

Parent–Child Relationship

Children have some of their first experiences with internal affective states, including anger, fear, and happiness, in the context of their relationships with their parents. Furthermore, the quality and intensity of children's emotional experiences are affected by the quality of their relationships with their parents. Parents may be highly influential in young children's regulation of their own affect. For example, in parents with their own emotional difficulties, such as depression or anxiety, repeated experiences with enduring intense negative emotion can be particularly challenging for young children as they attempt to regulate their own emotions. Dimensions of parenting pertaining to emotionality (e.g., acceptance, emotional availability, sensitivity, and attachment) have been found to have substantial implications for children's emotional development. Greater acceptance and responsiveness to children's needs by parents is strongly predictive of greater sociability, self-regulation, and prosocial behavior in children, whereas parental behaviors that are more indicative of a lack of responsiveness and availability are linked with poorer outcomes in children, such as increased aggressiveness and social withdrawal.

The effects of parental behavior on children's emotional adjustment are more than a matter of just the behaviors that parents direct toward their children or even the emotional intensity of parent–child interactions or parenting. Rather, they reflect the underlying emotional quality of the relationship between parents and their children. Thus, the ways parents and children interact together are influenced by the emotional bond or attachment that has formed between the parent and child.

John Bowlby and Mary Ainsworth contend that infants' patterns of interactions with their parents develop into patterns of attachment style. In infancy and early childhood, attachment security is assessed based on Ainsworth's Strange Situation task, which consists of a sequence of brief contexts for observing the children's functioning in relation to the parent's presence, absence, and return. Children's attachment securities are classified to distinguish parent–child

relationships in terms of the child's effectiveness in deriving security from the parent in these various contexts, as well as the parent's effectiveness in providing security.

Parents who are sensitive, responsive, and consistent in meeting their infants' needs over time are more likely to have infants who develop a secure emotional attachment. Children with secure attachments are able to optimally use their parents as a secure base and as support in the context of the attachment relationship. Thus, the child demonstrates a coherent strategy for using the parent as a source of security. For example, after the parent returns from the separation, the child seeks to reconnect with the parent and is able to recover easily from any emotional distress caused by the separation, returning to happily playing. Children who have a secure attachment feel confident that their parents are there for them when they need them and are able to successfully reach their parents in times of physical, social, emotional, or other needs and be comforted.

When parents are unable to consistently meet their infants' needs, then an insecure emotional attachment is more likely to develop. Children who are not securely attached may not feel safe or comfortable seeking help from their parents because they have not successfully had their needs met in the past. Instead, they may try to hide their emotions or needs. This is typical with children with avoidant attachments. Upon the parents' return, avoidantly attached children conspicuously avoid proximity or contact with their parents, and they are not responsive to parental attempts at interaction (e.g., they may look away, turn away, and fail to proactively initiate interaction). They are more fussy and distressed during the separation and yet have more difficulty in arousal control upon reunion. Parents of avoidantly attached children are more tense, irritable, and avoidant of close physical contact and overstimulation, thus fostering less confidence in their children about the parents as reliable sources of comfort and security.

Children can also demonstrate anxious or resistant insecurity in their attachment relationships with their parents. They too are unable to effectively use their parents as a source of security in times of stress, and their strategies reflect extreme dependence. Anxious or resistant children are more clingy initially, and when insecurely attached children do attempt to seek out their parents for help, they mix anger (e.g., struggling while being held, holding their bodies rigidly, hitting, or pushing away) with excessive contact and proximity seeking. They are not comforted by their parents' attempts at consoling and have considerable difficulty settling down and returning to a more regulated state.

Bowlby and Ainsworth argue that these early attachment patterns form the basis for all other relationships that children will develop. Children use their early attachment relationships and emotional experiences with their parents as the interpretative lenses through which to view, interpret, and respond to all other relationships and interactions. In other words, they develop internal representations or schemas about how relationships work based on their relationships with their parents. Levels of parent–child attachment security provide a global sense of emotional security that influences children in many broad domains. Furthermore, children who have successfully had their emotional needs met early on are better at regulating their own emotions as they develop and thus have more success in modulating their outward expressions of emotions, such as their behavior and other communication, because they can better modulate their internal states.

Marital Relationship

The parents' marital relationship is another key factor contributing to children's emotional development. The quality of the marital relationship and the ways marital conflict are handled set the emotional climate for the family and the child. Marital conflict has both positive and negative elements however, and depending upon how it is handled, it can have either positive or negative effects on children's emotional development. Heightened levels of negative conflict can influence children directly and indirectly.

Increased conflict can directly influence children, leaving them feeling vulnerable and emotionally insecure about the stability of their family. E. Mark Cummings and Patrick Davies proposed an emotional security hypothesis, which extends the ideas from attachment theory and suggests that children can develop a sense of emotional security not just to the parent–child relationship but also to the marital relationship. From their hypothesis, children react to the meaning of conflict itself, rather than just the presence of conflict, and the more they are exposed to it, the more sensitized they become to it over time. Witnessing more negative hostile and threatening forms of conflict is particularly detrimental to children's sense of security about the marital relationship, whereas

more positive expressions including humor, support, and affection promote greater feelings of security.

Children as young as 10 months of age respond to marital conflict with distress and act markedly different to conflict than to marital harmony. Children's reactions to expressions of anger and affection in the home change over time. In middle childhood, compared with toddlerhood, children are less likely to express their emotions overtly (e.g., by yelling, crying, laughing) during angry interparental exchanges. As children develop, they are also increasingly more likely to intervene in marital conflict scenarios.

Increased marital conflict can also influence children indirectly, through negative changes in parenting. Marital conflict is associated with inconsistent child rearing and disciplinary problems, increased parental negativity, intrusive control, and low levels of parental warmth and responsiveness. Conflict in the marital relationship may translate directly into angry interactions with children or may drain parents of the energy and emotional resources necessary to effectively parent their children.

The negative effects of conflict on the parent–child relationship and children's emotional development are well established. Increases in marital conflict during infancy are predictive of insecure parent–child attachment. Children's relationships with their parents may also change because of the negative effects on their sense of trust or positive regard for their parents caused by watching them behave in mean or hostile ways toward each other.

Normal and Abnormal Emotional Development

Early generations of research on atypical emotional development documented correlations between childhood factors and later developmental outcomes. For example, describing correlations between high levels of marital conflict or negative parenting and child depression or aggression. A second generation of research is moving beyond mere correlations and seeking to understand the processes underlying children's emotional development. This is exemplified in the field of developmental psychopathology, which has underscored the importance of understanding the processes underlying normal and abnormal development.

The developmental psychopathology approach calls attention to the importance of understanding the multiplicity of individual, biological, social, familial, and other processes underlying the development of childhood problems. For example, multiple emotional, cognitive, behavioral, and physiological processes have been implicated as mediators between marital conflict and children's adjustment. The developmental psychopathology approach seeks to study children over time to provide an understanding of the developmental processes and pathways that precede and account for the development of clinical disorders in children. The developmental psychopathology approach also advocates simultaneously examining both abnormal and normal development and risk and resilience to provide a more accurate, appropriately complex and comprehensive picture of the processes that account for the risk and emergence of psychopathology in children. Psychopathology is viewed as reflecting deviations from normative patterns over time. Inherent in this concept are the notions of *multifinality*, in which the same pathways lead to different outcomes, and *equifinality*, in which more than one pathway leads to the same outcome. The approach assumes that change is possible at any point along the pathway toward abnormality or normality, although change is constrained by prior adaptation. Furthermore, emphasis is placed on the significance of context for understanding developmental patterns: What may be dysfunctional or harmful in one context may be adaptive in another.

The developmental psychopathology approach is particularly adept in understanding why some children who experience negative environmental circumstances growing up (e.g., high marital conflict, poor parenting) develop emotional difficulties, whereas others do not. For example, in examining children of depressed parents, the pertinence of such an approach is highlighted. Children of depressed parents are at heightened risk for a full range of adjustment problems, including emotional difficulties such as sadness, depression, and anxiety. However, not all children with depressed parents develop problems, and not all children develop problems at the same point in their development. This provides evidence that other environmental factors must be considered. The developmental psychopathology approach leads researchers to examine a multiplicity of factors at multiple levels to understand the processes that modify children's adjustment. For example, children may be affected by parental depression through direct exposure, altered parent–child interactions and attachment, and increased marital conflict and family discord.

Thus, it is evident there is a need for a complex flexible theoretical model that can incorporate these diverse findings and yield a viable explanation of the multiple potential pathways of development. It is not as simple as just a genetic predisposition causing emotional problems, and it is not just the presence or absence of certain factors that can lead to normal or abnormal emotional development; rather, it is the way in which these factors transpire and interact that helps account for children's adjustment at any given time.

Tina D. Du Rocher Schudlich

See also Attachment; Family Influences; Parenting; Parenting Styles; Risk Factors and Development; Social Development

Further Readings

Ainsworth, M. D. S., Blehar, M., Waters, E., & Wall, S. (1978). *Patterns of attachment.* Hillsdale, NJ: Lawrence Erlbaum.

Bowlby, J. (1969) *Attachment and loss: Vol 1. Attachment.* New York: Basic Books.

Buss, A. H., & Plomin, R. (1984). *Temperament: Early developing personality traits.* Hillsdale, NJ: Lawrence Erlbaum.

Cummings, E. M., & Davies, P. T. (1994). *Children and marital conflict: The impact of family dispute and resolution.* New York: Guilford Press.

Cummings, E. M., Davies, P. T., & Campbell, S. B. (2000). *Developmental psychopathology and family processes: Theory, research, and clinical implications.* New York: Guilford Press.

Fox, N. A. (1994). The development of emotion regulation: Biological and behavioral considerations. *Monographs of the Society for Research in Child Development*, 59(2–3, Serial No. 240), 243–246.

Kopp, C. B. (1989). Regulation of distress and negative emotions: A developmental view. *Developmental Psychology*, 25, 343–354.

Emotional Intelligence

Presently there is no one definition of *emotional intelligence (EI)* that is universally accepted. Psychologists have approached EI from different vantage points. Some researchers see it as the interplay of cognitive ability with emotional knowledge and regulation and have utilized self-report measures to assess the construct (e.g., Reuven Bar-On, Daniel Goleman). Other researchers have viewed EI as a distinct intelligence and have measured it in the same way as traditional IQ (e.g., John Mayer, Peter Salovey, and David Caruso). *EI* is a term that has been defined and revised by theorists and researchers in the field of psychology. A useful definition has been proposed by Mayer and Salovey, who have noted that EI is the ability to perceive accurately; appraise and express emotions; access and/or generate feelings when facilitating thought; understand emotions and emotional knowledge; and regulate feelings to promote cognitive and emotional growth.

EI is most relevant because it can be a powerful predictor of how successful a person can be in life, and it has been postulated that if children are appropriately taught, EI can be increased. In this entry, definitions of EI will be presented, ways of measuring the construct will be explained, and significant research findings will be discussed.

At Stanford University in the 1960s, a researcher made the following proposition to 4-year-old children. "I have to do an errand; if you can wait until I come back, you can have two marshmallows for a treat. If you cannot wait, you can have only one. But you can have it right now." Then the researcher left.

Following his departure, some children grabbed the marshmallow as soon as he walked out the door. Others lasted for a few minutes, but eventually the temptation was too much. Those with emotional self-control were determined to wait. These youngsters covered their eyes, or put their heads down. Others turned around in their chairs or distracted themselves by playing games or singing. Eventually the experimenter returned and gave the patient children their extra treat.

Twelve to 14 years later, these children were evaluated as teenagers. Those 4-year-olds who were able to resist temptation were now, as adolescents, viewed by their parents as more socially competent, personally effective, self-assertive, and better able to cope with the frustrations of life. They tended to be more confident, trustworthy, and dependable. They were better at taking initiative and developing projects. And, they were still better able to delay gratification to pursue their goals. In addition, those children who, at age 4, waited for the researcher to return had dramatically higher SAT scores as high school seniors. The third of children who grabbed at the marshmallow as preschoolers had an average verbal score of 524 and a quantitative score of 528. The third who waited the longest had average SAT scores of 610 (verbal) and 652

(quantitative)—a 210 difference in total score. This study, conducted by psychologist Walter Mischel, demonstrates both the short-term and future advantage of restraining emotions and delaying impulses. This is the essence of emotional self-regulation—one domain of emotional intelligence. Individuals who are better at emotional self-regulation are more primed for success—whether it is pursuing a career, building a business, or establishing relationships.

What Is Emotional Intelligence?

The term *emotional intelligence* entered the American lexicon in 1995 when Daniel Goleman published his book *Emotional Intelligence: Why It Can Matter More Than IQ.* It became a bigger buzzword, when Nancy Gibbs wrote an October 2, 1995, *Time* magazine cover story, titled "The EQ Factor." Her emphasis was that new brain research suggested that emotions, not IQ, may be the true measure of human intelligence.

Though Goleman received the most attention for his groundbreaking work, Peter Salovey and John Mayer originally used the term *emotional intelligence* in 1990 and later refined their definition in 1997. From their theoretical perspective, EI refers specifically to the cooperative combination of intelligence and emotion. The authors label their model as a four-branch ability model and divide the abilities and skills of EI into four areas: the abilities to (1) perceive emotions, (2) use emotions to facilitate thought, (3) understand emotions, and (4) manage emotions.

Branch 1 reflects the perception of emotions and involves the capacity to recognize emotions in another person's facial and postural expressions. Branch 2, facilitation, involves the capacity of emotions to assist thinking. Branch 3, the understanding of emotion, reflects the capacity to analyze emotions, appreciate their probable trends over time, and understand their outcomes. Branch 4 reflects the management of emotions, which involves the rest of personality. Emotions are managed in the context of the individual's goals, self-knowledge, and social awareness.

A second definition of EI was postulated by Goleman based on the work of Salovey. He defined EI as the capacity for recognizing one's own feelings and those of others, for motivating oneself, and for managing emotions well in oneself and in one's relationships. He sees these abilities as distinct from, but complementary to, cognitive intelligence traditionally measured by IQ tests.

He noted that these two types of intelligence—cognitive and emotional—relate to different areas of the brain. The intellect is based on the workings of the neocortex, the more evolved part of the brain. The emotional centers are in the more primitive subcortex. EI involves the interplay of the emotional centers in tandem with the intellectual centers.

In developing his definition, Goleman expanded on the work of Howard Gardner, who developed the theory of multiple intelligences. Two of these multiple intelligences were personal in nature: interpersonal intelligence (social skills) and intrapersonal intelligence (self-knowledge). In 1998, in the text *Working With Emotional Intelligence*, Goleman modified his focus in order to understand how these talents work in life. He began to label these as emotional and social competencies, which he described as learned capabilities based on emotional intelligence that results in outstanding performance at work. These include the following:

Self-awareness (knowing one's internal states, preferences, resources, and intuitions). This is self-awareness or recognizing a feeling as it happens. This is considered to be the keystone of EI. These competencies include emotional awareness, accurate self-assessment, and self-confidence.

Self-regulation (managing one's internal states, impulses, and resources). This relates to the ability to soothe oneself; that is, to shake off depression, irritation, stress, and anxiety. This is the fundamental skill of emotional self-regulation. He sees these competencies as including self-control, trustworthiness, conscientiousness, adaptability, and innovation.

Motivation (emotional tendencies to guide or facilitate reaching goals). This relates to the ability to self-motivate oneself. These competencies include achievement drive, commitment, initiative, and optimism.

Empathy (awareness of other's feelings, needs, and concerns). These competencies include understanding others, helping others develop, developing a service orientation, leveraging diversity, and maintaining political awareness.

Social skills (adeptness at inducing desirable responses in others). This requires handling emotions in relationships well and accurately reading social situations and networks; interacting smoothly; using these skills to persuade and lead, negotiate, and settle disputes for cooperation and teamwork. These competencies

include wielding influence, communicating effectively, managing conflict, providing leadership, serving as a change catalyst, building bonds, collaborating and cooperating, and creating team capabilities.

The first three—self-awareness, self-regulation, and motivation—are considered personal competencies; whereas the last two—empathy and social skills—are considered social competencies.

The third major contributor to the definition of the construct of EI is Bar-On. He refers to this construct as emotional-social intelligence (ESI). Bar-On views ESI as a cross section of interrelated emotional and social competencies, skills, and facilitators that determine how effectively one understands and expresses oneself, understands others and relates with them, and copes with daily demands. These include the following five components:

1. The ability to recognize, understand, and express emotions and feelings

2. The ability to understand how others feel and relate with them

3. The ability to manage and control emotions

4. The ability to manage change, adapt, and solve problems of a personal and interpersonal nature

5. The ability to generate positive affect and be self-motivated

How Is Emotional Intelligence Measured?

Measurement of EI has involved two major approaches, according to Gerald Matthews, Moshe Zeidner, and Richard Roberts. The first measurement approach has aligned with Mayer, Salovey, and Caruso's model of EI. To be considered a standard intelligence, Mayer, Salovey, and Caruso suggest that EI must (a) be operationalized as a mental ability, (b) meet correlational criteria that indicate it as a unitary ability that represents a new kind of performance relative to earlier measures of intelligence and personality dispositions, (c) exhibit growth with age, and (d) predict outcomes of importance.

These three authors have developed the Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT), which meets all four criteria. This test has eight tasks: two to measure each of the four branches of EI. The MSCEIT is considered an ability test, similar

in nature to a more traditional IQ test. The examinee is asked to do a task and receives a score. For example, for Branch 1, Perceiving Emotions, in the subtest Faces, the examinee is asked to identify the emotions in faces. The authors conclude that the MSCEIT is a convenient-to-administer test that is highly reliable at the total score area and branch levels and provides a valid measure of EI.

The second measurement approach has utilized self-report measures to operationalize the models of EI presented by Goleman and Bar-On. Several measurements have been developed based on Goleman's model. One of these is the Emotional Competency Inventory, developed by Richard Boyatzis, Goleman, and Hay/McBer (now the HayGroup), which is a multirater (360-degree) instrument that provides self, manager, direct report, and peer ratings on a series of behavioral indicators of EI. It measures 20 competencies, and the rater is asked to describe him- or herself or the other person on a scale from 1 (the behavior is only slightly characteristic of the person) to 7 (the behavior is very characteristic of the person) for each item. These items are then composed into ratings for each of the competencies.

Bar-On developed a self-report measure of emotionally and socially intelligent behavior called the Emotional Quotient Inventory (EQ-i). It was the first measure on this construct published by a psychological test publisher, the first to be peer reviewed, and the most widely used measure of its kind. It contains 133 items in the form of short sentences and employs a 5-point response scale. It has five composite scales and 15 subscales. The composite scales are (1) intrapersonal, (2) interpersonal, (3) stress management, (4) adaptability, and (5) general mood.

What Are the Significant Research Findings?

Views of EI differ according to the purpose of the researchers (e.g., studying EI as a standard intelligence or as the capacity for processing and regulating emotions). Consequently, there is no consensus definition of EI. A more precise definition should be expected to emerge as researchers test the assumptions behind the different models of EI. Notably, more research is needed to delineate EI from constructs such as temperament and personality. Programmatic research by Mayer, Caruso, Salovey, and colleagues is beginning to delineate how EI converges and diverges from standard intelligence.

Research utilizing Mayer, Salovey, and Caruso's model suggests that there is a mild to medium overlap between general EI and verbal intelligence (i.e., crystallized intelligence). In contrast, results of a meta-analysis of 10 studies suggest that no more than 4% of the variance of the EQ-i can be explained by cognitive intelligence. This suggests that ESI and cognitive intelligence are not strongly related and are most likely separate constructs. In addition, there is neurological evidence suggesting that neural centers governing ESI and those regulating cognitive intelligence are located in different areas of the brain. Specifically, it has been hypothesized that the ventromedial prefrontal cortex appears to be related to basic dimensions of ESI, whereas the dorsolateral prefrontal cortex is thought to govern significant aspects of cognitive intelligence.

No differences have been revealed between males and females regarding overall ESI. However, there are a few statistically significant though small differences between the sexes for a few of the factors. Females appear to have stronger interpersonal skills than males, but males have a higher intrapersonal capacity, are better at managing emotions, and are more adaptable. The Bar-On model reveals that women are more aware of emotions, relate better interpersonally, demonstrate more empathy, and are considered more socially responsible than men. In contrast, men appear to be more self-reliant, cope better with stress, be more flexible, solve problems better, be more optimistic, and have better self-regard.

In terms of age, results from the EQ-i revealed that older groups scored significantly higher than the younger groups on most scales, and individuals in their late 40s have obtained the highest mean scores. The results seem to suggest that as a people get older, they become more socially and emotionally intelligent.

Current research has been exploring the relationships between EI, behavior, and performance (e.g., school grades, self-actualization, stress management, wellness, leadership, and occupational performance). Some noteworthy studies are highlighted here. In a path analysis study, James Parker and colleagues found that there was a .41 correlation between ESI and scholastic performance among high school students, indicating a moderate statistically significant correlation. Results seem to indicate that at least 17% of academic performance is a function of ESI in addition to cognitive intelligence.

A series of studies using large samples from the Netherlands, Israel, and North America evaluated the relationship between self-actualization and EI. Results indicated that emotional-social intelligence more so than cognitive intelligence influences one's ability to do one's best. Bar-On notes that a high IQ does not guarantee that one will actualize their potential, but a high EQ, more so than a high IQ, gives a person a better chance. This bolsters the notion that adults who eventually reach their goals are those who, as children, had the self-discipline to wait for the extra marshmallow.

Conclusion

Emotional life requires a unique set of competencies and as surely as any academic or cognitive domain can be managed with lesser or greater skill. As noted by Goleman, how adept a person is at EI is necessary to understanding why one person excels in life while another with equal cognitive skills falls short. Emotional aptitude may be conceptualized as a meta-ability that determines how well a person uses their raw intellect to navigate the rivers and shoals of life.

Philip J. Lazarus and Nick Benson

See also Crystallized Intelligence; Fluid Intelligence; Intelligence and Intellectual Development; Intelligence Quotient (IQ); Intelligence Tests

Further Readings

Bar-On, R. (1997). *The Emotional Quotient Inventory (EQ-i): A test of emotional intelligence.* Toronto, Ontario, Canada: Multi-Health Systems.

Bar-On, R. (2000). Emotional and social intelligence: Insights from the Emotional Quotient Inventory (EQ-i). In R. Bar-On & J. D. A. Parker (Eds.), *Handbook of emotional intelligence* (pp. 363–388). San Francisco: Jossey-Bass.

Bar-On, R. (2006). The Bar-On model of emotional-social intelligence (ESI). *Psicothema, 18*(Suppl.), 13–25.

Boyatzis, R. E., Goleman, D., & HayGroup. (2001). *The Emotional Competence Inventory (ECI).* Boston: HayGroup.

Brackett, M. A., & Salovey, P. (2004). Measuring emotional intelligence with the Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT). In Glenn Geher (Ed.), *Measuring emotional intelligence: Common ground and controversy* (pp. 181–196). Hauppauge, NY: Nova Science.

Gibbs, N. (1995, October 2). The EQ factor. *Time,* 60–68.

Goleman, D. (1995). *Emotional intelligence.* New York: Bantam Books.

Goleman, D. (1998). *Working with emotional intelligence.* New York: Bantam Books.

Matthew, G., Zeidner, M., & Roberts, R. D. (2002). *Emotional intelligence: Science and myth.* Cambridge: MIT Press.

Mayer, J. D., & Salovey, P. (1997). What is emotional intelligence? In P. Salovey & D. Sluyter (Eds.), *Emotional development and emotional intelligence: Implications for educators* (pp. 3–31). New York: Basic Books.

Mayer, J. D., Salovey, P., & Caruso, D. R. (2002). *Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT).* Toronto, Ontario, Canada: Multi-Health Systems.

Mayer, J. D., Salovey, P., & Caruso, D. R. (2004). Emotional intelligence: Theory, findings and implications. *Psychological Inquiry*, *15*, 197–215.

Parker, J. D. A., Creque, R. E., Barnhart, D. L., Harris, J. I., Majeski, S. A., Wood, L., et al. (2004). Academic achievement in high school: Does emotional intelligence matter? *Personality and Individual Differences*, *37*, 1321–1330.

Salovey, P., & Mayer, J. D. (1990). Emotional intelligence. *Imagination, Cognition, and Personality*, *9*, 185–211.

Stys, Y., & Brown, S. L. (2004). *A review of the emotional intelligence literature and implications for corrections.* Ottawa, Ontario, Canada: Correctional Service of Canada, Research Branch.

EMOTION AND MEMORY

People's memories of past events possess different qualitative characteristics, including overall vividness. Certain events are recalled with high clarity and a great degree of detail, whereas other events are barely remembered. Research on the effects of emotion on memory has shown that emotion can play an important role in determining such differences in recall. Positive, negative, and neutral emotions can influence whether an event or information will be richly and vividly recollected. Investigations of emotion and memory have included studies of emotion and general autobiographical memories, eyewitness testimony, flashbulb memories, and memories for traumatic events. This entry examines each of these topics.

Emotion and General Autobiographical Memories

Events that are pleasant are often processed at a deeper level and, thus, recalled more accurately and faster than events that are unpleasant. Due to more vivid imagery associated with pleasant items, memory for positive information is more accessible, durable, and frequent than is memory for negative information. Margaret Matlin proposed that this phenomenon is part of the *Pollyanna Principle.* Another aspect of the Pollyanna Principle is the *fading affect bias*, which states that unpleasant memories fade faster than pleasant memories, because the emotion associated with unpleasant events weakens in intensity more than the emotion associated with pleasant events. Thus, overall, people have the tendency to focus on positive life experiences and are motivated to view their life events as relatively pleasant.

However, the Pollyanna Principle does not seem to apply to people with depressive tendencies. Individuals who are depressed tend to focus on negative events, and as a result, unpleasant emotions do not fade more quickly than pleasant emotions. With a profound sense of hopelessness, depressed individuals' autobiographical memories are biased toward unpleasantness. These individuals often recall *mood-congruent* material; that is, they have the tendency to recall more negative than positive material because negative material is congruent with their current mood. Depression is linked to less specific retrieval of positive memories. For example, investigations of patients who have recently attempted suicide reveal that their recent autobiographical memories consist of mostly negative episodes. Furthermore, there is a delay in these patients' ability to retrieve positive memories. Individuals who are at risk for depressive moods possess automatic negative and depressive biases, while also attempting to suppress their negative tendencies to inhibit the influence of depressive biases. As a result of this conflict, they often become relatively uncertain about the meaning of ambiguous information, and their memory for such information is tainted.

Emotion and Eyewitness Testimony

Individuals who experience a history of negative or traumatic events, such as a history of sexual or physical abuse, have reduced memory specificity. As a regulating strategy, these individuals routinely retrieve autobiographical memories in a less specific way to avoid being confronted with past painful memories. This strategy may be perceived as advantageous and protective in the short term. However, as Dirk Hermans and his colleagues suggested, in the long term, a *cognitive avoidance* strategy may turn out to be

maladaptive, due to using an avoidant coping style to deal with feelings, thoughts, and problems.

Overall, about one in three people who have experienced a traumatic event report having memory difficulties retrieving details about the event, probably because they were so emotionally distracted that the normal memory encoding processes were interrupted at the time the event occurred. Researchers in the area of eyewitness testimony have indicated that similar to other types of traumatic events, witnessing a crime involves intense emotion, and such experience is capable of disrupting memory encoding and recall. The strong emotions that are elicited in a crime situation tend to draw attention to the central, salient themes of the sequence of the event, and thus, memory for the core elements is facilitated by the experience. On the other hand, peripheral details of the events are often impaired because attention is drawn to critical themes, leaving little capacity to process specific details. The phenomenon is known as the *Easterbrook hypothesis*, which states that in a traumatic condition, such as a crime situation, high emotional arousal narrows one's range of focus on the important aspects of the events, and as a result, memory for noncentral, irrelevant details is decreased or even distorted.

For example, when a bank robbery was depicted in a film in Elizabeth Loftus's research, participants who viewed a traumatic version, in which a boy was abruptly shot in the face, later recalled much less detail from the beginning of the film (e.g., the number on the boy's jersey), compared to participants who saw a nontraumatic version of the film. This finding reveals that a high degree of anxiety and emotional intensity will narrow attentional focus, drain available processing resources, and reduce memory ability. Directly related to the Easterbrook hypothesis is the *weapon focusing hypothesis.* When a weapon is used in a crime, witnesses often have clear memories of the weapon involved but hindered memories of other aspects of the crime, including central elements like the criminal's face. This phenomenon is especially applicable to people who have a high state of anxiety when seeing a weapon. A heightened sense of anxiety can trigger panic that leads to distorted memory, because different aspects of the event are not well encoded into memory.

Emotion and Flashbulb Memories

Another intriguing phenomenon in the study of emotion and memory was first referred to as *flashbulb memory* by Roger Brown and James Kulik. These investigators proposed that individuals' memory for a shocking and novel event, like the assassination of John F. Kennedy, often seems to be preserved in a photographic form that contains many circumstantial details that are resistant to forgetting. Brown and Kulik suggested that this type of memory is processed through a special brain mechanism called *Now Print*, in which the brain's flashbulb, like the camera's flashbulb, freezes the moment when the unusual event occurs, and, consequently, memory for the event is stored with piercing clarity and precise details. However, a key element that determines whether the Now Print mechanism is activated is the consequentiality of the event. A surprising event that has high personal involvement and significance is more likely to trigger a flashbulb memory. For instance, according to research findings, more African Americans than European Americans reported possessing flashbulb memories of the assassination of Martin Luther King, Jr., and Californians showed more flashbulb memories that were accurate and detailed for the 1989 earthquake in Loma Prieta (near San Francisco) than Atlanta residents who heard about the event from the news.

Other investigators have argued against Brown and Kulik's view that flashbulb memories are formed through a special brain mechanism, because research has revealed that flashbulb memories are prone to forgetting over time, just like memories for ordinary events. People's recollections of shocking, highly emotion-driven public events, such as the *Challenger* explosion, the bombing of Iraq in 1991, and the verdict of the O. J. Simpson trial, often contain distortions and can be remarkably inaccurate. As time passes, people forget or become confused about the circumstances in which they first heard about the event, and, thus, their memories for the event can show substantial degrees of decaying and fading. Nevertheless, false flashbulb memory recollections are often accompanied by high degrees of confidence by individuals. In other words, despite inaccurate recall, individuals tend to be very confident about their false memories being correct, leading to a mismatch between subjective feelings of confidence and objective accuracy of recollections.

Although flashbulb memories can fade and produce reconstructive errors over time, they are sometimes more durable and accurately remembered than many memories of ordinary events. These vivid flashbulb memories are likely to be strengthened through

repetitive discussion and rehearsal. People talk and think about the surprising event repeatedly for an extended period of time following the event, resulting in an increased memorability for the event. Overall, intense emotion and personal significance associated with an event lead to more accurate and vivid remembering of the event because these elements make an event distinctive and different from mundane occurrences.

Emotion and Memories for Traumatic Events

It is a common belief that memory for emotionally traumatic events is accurate, vivid, and detailed. Although it is often true that memory for emotionally laden traumatic events is fairly well preserved, it is not an exception to find striking errors and distortions in people's recollections of past traumatic life events. For example, it has been shown that recollections of children who have been through certain traumatic events, such as being kidnapped at gunpoint or being attacked by a sniper at school, often contain a significant number of inaccuracies, even though their memories appear to be vivid and full of details. Distortions of traumatic memories can be the result of perceptual errors due to stress experienced during the shocking episode. Seeing a mustache on the attacker's face when the attacker does not have facial hair is an example of a perceptual error caused by a high level of emotion during the event. Traumatic memories can also change over time and, thus, contain errors because of retrospective biases. Memory of an event can be later reconstructed in a way to fit a person's current emotional state after some time has passed. For example, people with a reduced level of posttraumatic stress tend to recall an original traumatic event as less threatening than their original recall of the event compared with those with a heightened level of posttraumatic stress.

Children are more vulnerable to the influence of misinformation and more apt to confuse the sources of their knowledge than are adults, resulting in mistakenly incorporating inaccurate information into their memory or failing to distinguish between an imagined and an actual event. However, inaccurate memories of traumatic events have also been demonstrated in adults who survived some brutally traumatic events like Nazi concentration camps. Thus, even extreme trauma cannot ward off possible memory erosions over time. Similarly, combat flashbacks, often reported by war veterans, can be as visually intense as if the individuals are reliving moments of the actual combat episodes. And yet, under close examination, this type of flashback is often a combination of memories of real events and constructions of imagined events. People who are easily hypnotized and have an inclination to engage in imaginative and fantasy-based activities are most likely to make claims of having flashbacks of traumatic past experiences. The authenticity and truthfulness of such flashbacks, however, have been highly questioned by researchers.

Further Research

Emotion seems to play a major role in how people's memories are initially constructed and later reconstructed during the retrieval process. High levels of emotion may enhance some aspects of memories while impairing other aspects of recollections. Investigators have used various paradigms when studying the effect of emotion on memory. Further research on this topic will continue to shed light on the role that people's emotional lives play in their autobiographical memories and on the mechanisms by which emotions exact their toll on memory.

Lin-Miao L. Agler and Karen M. Zabrucky

See also Episodic Memory; Flashbulb Memories, The Nature of; Long-Term Memory; Memory

Further Readings

Hermans, D., Defranc, A., Raes, F., Williams, J. M. G., & Eelen, P. (2005). Reduced autobiographical memory specificity as an avoidant coping style. *British Journal of Clinical Psychology*, *44*, 583–589.

Howes, M. B. (2007). *Human memory: Structures and images.* Thousand Oaks, CA: Sage.

Schacter, D. L. (1996). Emotional memories: When the past persists. In D. L. Schachter, *Searching for memory: The brain, the mind, and the past* (pp. 192–217). New York: Basic Books.

Schmolck, H., Buffalo, E. A., & Squire, L. R. (2000). Memory distortions develop over time: Recollections of the O. J. Simpson trial verdict after 15 and 32 months. *Psychological Science*, *11*(1), 39–45.

Swales, M. A., Williams, J. M. G., & Wood, P. (2001). Specificity of autobiographical memory and mood disturbance in adolescents. *Cognition and Emotion*, *15*(3), 321–331.

Walker, W. R., Skowronski, J. J., Gibbons, J. A., Vogl, R. J., & Thompson, C. P. (2003). On the emotions that accompany autobiographical memories: Dysphoria

disrupts the fading affect bias. *Cognition and Emotion*, *17*(5), 703–723.

Walker, W. R., Skowronski, J. J., Thompson, C. P. (2003). Life is pleasant—and memory helps to keep it that way! *Review of General Psychology*, *7*(2), 203–210.

Williams, J. M. G., & Broadbent, K. (1986). Autobiographical memory in suicide attempters. *Journal of Abnormal Psychology*, *95*(2), 144–149.

EMPATHY

The empathic teacher and the empathic student are important topics for consideration by educational psychologists. Contemporary approaches to empathy conceive of empathy as a social interaction between any two individuals with one individual experiencing the feelings of a second individual. Although there is some degree of correspondence between the affect of the observer and the affect of the observed, the affects are not identical. The process of empathy is currently acknowledged to be contingent on both cognitive and affective factors, the particular influence varying with the age and other attributes of the individual and with the situational context. The model proposed by Norma Feshbach emphasizes the cognitive ability to discriminate affective states in others, the more mature cognitive ability to assume the perspective and role of another person, and the affective ability to experience emotions in an appropriate manner. Martin Hoffman's developmental model also has three components—cognitive, affective, and motivational—and focuses on empathic responsiveness to distress in others as the motivation for altruistic behavior.

Definitional concerns, methodological problems, and theoretical controversies have characterized this area of study. Nevertheless, the status of empathy as an important variable meriting consideration and empirical study has dramatically changed during the past 30 years. Its relevance to a number of disciplines, including neuroscience, psychology, sociology, political science, education, medicine, and the arts, has no doubt contributed to the burgeoning interest and activity in the study of empathy.

Measurement

The task of assessing empathy remains a formidable problem. The multidimensionality of the construct and the internal properties of its components leave a wide hiatus between (a) the critical role afforded to empathy in the individual's development and social behavior and (b) its data base. Assessment procedures for children's empathy vary as a function of stimulus modality (e.g., stories, audiotapes, cartoons, paper-and-pencil questions, slides, and narration), as well as response modality (e.g., reflexive crying and self-reports). More recently, measures of physiological arousal, such as heart rate and electrodermal responses, are being used to assess empathy.

Adult measures tend to be paper-and-pencil, questionnaire-type instruments. However, recent measurement developments entail the analysis of facial musculature responses to others and neuronal responses, the latter pertaining to so-called mirror neurons that match the neural reactions of another person.

Origins

The answer to the ontogenetic pattern of empathic development is unresolved. Infant responsiveness to the crying of other infants has been noted in infants as young as 2 weeks. Studies also show that infants younger than 1 year of age can respond differentially to faces depicting different affective states. In the studies by Carolyn Zahn-Waxler and Marian Radke-Yarrow, very young children exhibited "empathy-like behaviors." And while there is a close theoretical and empirical relationship between empathy and prosocial behaviors, empathy is theoretically distinct from and should not be considered as synonymous with helping, caring, sharing, or any other positive social behavior. Generally, it can be said that a cluster of empathy-related skills appear very early in an individual and become more differentiated and purposeful with age. Individuals, at any age, vary in empathic responsiveness, the source of which may be biologically/temperamentally or situationally/environmentally determined.

Antecedents of Empathy

The literature relating parent attitudes and child-rearing practices to children's empathy as well as laboratory and field training studies indicates that empathy can be taught and that empathy can be learned. How parents rear their children influences the development of empathic understanding and behavior. Mothers who are responsive, nonpunitive, and nonauthoritarian and who manifest empathic and caring behaviors enhance the development of empathy. Conversely,

parent practices that are negatively related to the development of empathy include use of threats and physical punishment, inconsistency, and reliance on extrinsic rewards. The strength of these parent–child interrelationships is much greater for mothers and their daughters than for fathers with either their sons or daughters, data that are consistent with overall findings that females are modestly, but consistently, higher in empathy than males.

Functions of Empathy

Empathy appears to play an important mediating role in a wide range of cognitive, affective, and social behaviors. These may include greater social understanding, greater emotional competence, heightened compassion and caring, greater regulation of aggression and other antisocial behaviors, and greater prosocial and moral behaviors. Studies relating empathy to such prosocial behaviors as cooperation, sharing, donating, and other altruistic acts have generally yielded positive findings, especially in adults, as have studies relating empathy to reduced social prejudice.

Because of its multifaceted structure encompassing cognitive and affective dimensions, empathy is linked to other important domains of behavior, such as communication and academic achievement. There are a number of studies showing associations between empathic understanding or the training of empathic skill and academic performance; for example, the Bonner and Aspy study relating empathetic understanding and grade point averages, the Feshbach and Feshbach training studies, the Feshbach and Feshbach longitudinal study of the relationship of empathy to reading and spelling skills in elementary school-age children, and the Kohn program on "caring school communities." Empathy should be distinguished from emotional intelligence, a concept that has received considerable attention in recent years. However, empathy is proposed to be a component of emotional intelligence.

Teacher Empathy

The salience of empathy to the educational process is especially reflected in the extensive literature on teacher empathy that was initially fueled by Halsey Rogers's approach to the therapeutic process and human development. For Rogers, empathy entails an understanding of the perspective and feelings of the client. The assumption underlying a focus on teacher empathy is the expectation that empathic communication by the teacher will result in students experiencing greater acceptance and in students developing more positive attitudes toward themselves and toward schooling.

The thrust of many of the teacher empathy studies has been on instructional techniques for enhancing teacher empathy. Most of the instructional programs entail complex, multimodal methods with varied effects, making it difficult to link teacher empathy to student outcomes. However, a few studies linking teacher empathy to student learning and behavior have found that teacher empathy is related to positive student outcomes.

Empathy in the Classroom

Today, there are a number of programs that directly or indirectly use empathy as an adjunct tool in teaching children. Earlier efforts with regard to empathy and education focused on enhancing empathy in teachers as a component of teacher competence by training them in human relations and social understanding. The focus of more current research is to show the relationship between teacher empathy and student behavior.

Future Research

Empathy appears to be a significant variable in human development and social interaction. Its complexity and internal representation, while challenging, do not preclude consideration and investigation. Promising research developments, such as the discovery of neural correlates, may help resolve some of the ambiguity in the study of empathy. Although more research is required, the process of empathy appears especially germane to the field of education, particularly educational psychology.

Norma D. Feshbach

See also Emotional Development; Emotional Intelligence; Social Development

Further Readings

Carkhuff, R. R. (1969). *Helping in human relations* (Vols. 1 & 2). New York: Holt, Rinehart & Winston.
Eisenberg, N., Fabes, R., Schaller, M., Carlo, G., & Miller, P. (1991). The relations of parental characteristics and

practices to children's vicarious emotional responding. *Child Development*, *62*, 1393–1408.

Feshbach, N. D. (1975). Empathy in children: Some theoretical and empirical considerations. *Counseling Psychologist*, *5*, 25–39.

Feshbach, N. D., & Feshbach, S. (1982). Empathy training and the regulation of aggression: Potentialities and limitations. *Academic Psychology Bulletin*, *4*, 399–413.

Feshbach, N. D., & Feshbach, S. (2003). Empathy in education: An elixir but no panacea. *International Journal of Psychology*, *10*, 2763–2769.

Hoffman, M. L. (1982). Developmental prosocial motivation: Empathy and guilt. In N. Eisenberg (Ed.), *The development of prosocial behavior* (pp. 218–231). New York: Academic Press.

Kohn, A. (1991). Caring kids: The role of schools. *Phi Delta Kappan*, *72*, 496–506.

Rodman, G. I. (1977). Study of the relationship of teacher empathy for minority persons and inservice relations training. *Journal of Educational Research*, *70*, 205–210.

Upright, R. L. (2002). To tell a tale: The use of moral dilemmas to increase empathy in the elementary school child. *Early Childhood Education Journal*, *30*, 15–20.

Zahn-Waxler, C., Radke-Yarrow, M., & King, R. A. (1979). Child rearing and children's prosocial initiations toward victims of distress. *Child Development*, *50*, 319–330.

ENGLISH AS A SECOND LANGUAGE

English as a second language is English for nonnative speakers. Because it involves learning across the life span, it is sometimes considered to be a subfield of educational psychology. English for nonnative speakers is referred to by several acronyms: ESL (English as a second language), EFL (English as a foreign language), EAL (English as an additional language), and ELL (English language learning). There are also related acronyms, such as TESOL (teaching English to speakers of other languages). The history of ESL instruction and research will be covered in this entry, as will current practices and key issues in the field.

Introduction

English holds preferential status because it is studied as a second or additional language by more people than any other human language; it has become the global lingua franca—the language of commerce, science, technology, and other professional endeavors. It has been estimated that for every native speaker of English, there are four nonnative speakers who use it

as a second or other language. Braj Kachru described three principal contexts in which English is learned: the inner circle, the outer circle, and the expanding circle. In inner-circle countries, such as Britain, the United States, Australia, New Zealand, and Canada, the majority of speakers use English in all aspects of their lives. In the outer circle, in countries such as India, Pakistan, Hong Kong, and Singapore, English is important for historical reasons and, in many instances, is an official language, but it is not the mother tongue of many of the citizens. The expanding circle consists of countries where English was not important historically but where it is now widely used as an additional language; for example, Japan.

The theory and pedagogy of ESL draw on many disciplines, including linguistics, education, psychology, anthropology, and sociology. In the past several decades, an amalgam of these areas has developed as applied linguistics, a field that encompasses not only language learning and teaching, but translation, lexicography, and other applied language issues.

Communicative Competence

A central concept in ESL is *communicative competence*, a term first coined by Dell Hymes, a linguistic anthropologist. Communicative competence includes the skills and knowledge necessary to successfully convey meaning through language. Applied linguists Michael Canale and Merrill Swain developed a model of communicative competence that describes the components of language that are crucial to successful communication. These components are grammatical (linguistic), sociolinguistic, strategic, and discourse competence. *Grammatical competence* refers to the ability to manipulate language forms according to grammatical rules; a person demonstrates grammatical competence by using grammatical forms accurately, but he or she may not be able to articulate the rules (as is the case for most native speakers). *Sociolinguistic competence* is demonstrated when an individual uses language appropriately in different contexts. The status of participants, the purpose of the interaction, and local norms all affect what is considered to be appropriate; for example, one would use different language when speaking to a judge in court than one would when communicating the same message to a close friend. *Strategic competence* is the ability to use verbal and nonverbal communication strategies to rectify or avoid communication

breakdowns and to enhance communication. Several researchers have developed extensive lists of communication strategies, such as paraphrase, repetition, word coinage, clarification requests, and so forth. *Discourse competence* refers to the ability to link forms and meanings coherently in both written and spoken language. This includes using discourse markers, such as *first*, *next*, *then*, and *finally*, to help the listener follow a sequence of events, for example. Several researchers have developed elaborated versions of Canale and Swain's model, but the same skills and basic concepts are emphasized. Communicative competence is the ultimate goal of many ESL learners and teachers.

Historical Overview of Teaching English

Although the teaching and learning of second languages (L2s) has been a concern of educators for many centuries (e.g., John Amos Comenius wrote about his nouvelle méthode in the 1600s), this overview will be limited to the past 200 years. The oldest method of language teaching is grammar translation (sometimes called the classical method); this stemmed originally from the teaching of Latin in Europe but spread to other contexts. This method is still used today in some settings to teach English; in China, for example, the impact of the grammar translation method is still felt quite strongly. Grammar translation focuses on the translation of sentences and texts from the first language into the L2 and vice versa. In addition, explicit grammatical explanations are provided with examples, along with exercises that focus on each grammatical rule. Vocabulary instruction is based on the text that the students are translating; thus, vocabulary learning does not progress from simple to complex. There is no emphasis on speaking or listening, because the intent is to produce people who can read and write in the new language. An advantage of grammar translation is that it can be undertaken independently, without an instructor, or with an instructor who is not very proficient in the language to be learned. Grammar translation was favored initially by the educated classes; it was viewed not only as a method for learning to read classic works such as Shakespeare but also as a valuable mental exercise.

In 1880, François Gouin published a book outlining a new approach to language teaching called the *series method*. This method was a radical departure from grammar translation, founded on the premise that an L2 should be learned much like the first, focusing on everyday spoken language rather than literary texts. Gouin followed the language learning progress of his young nephew and developed the series method based on his observations. No translation is involved (hence, it belongs to a group of methods that are referred to as being "direct"); students are taught a series of connected sentences accompanied by appropriate actions or behaviors. For instance, to explain how to describe a girl lighting a stove, the students would learn sentences such as "She puts down the wood in front of the stove," "She crouches down in front of the stove," and "She opens the door of the stove." Another feature of the series method is the lack of grammatical explanation; students are expected to learn grammar implicitly through exposure in the series of sentences. Finally, the series method emphasized the importance of accurate pronunciation.

Around the same time that Gouin produced his method, Charles Berlitz and others introduced other direct methods to teach L2s. In the Berlitz approach, which is still in use today, students are taught in the L2 exclusively, and their lessons are carefully ordered from concrete, simple forms to increasingly more complex and abstract language. In the first lessons, students learn some nouns such as *pencil*, *pen*, *book*, *chair*, and so on—items that are in the immediate environment. They also learn a few questions and answers, such as "What is this?" and "This is a pen," to help generate additional language. As students progress, they learn more nouns, adjectives, verbs, and verb tenses. Grammatical points are introduced one at a time, in an inductive manner.

In the 1920s, there were several advocates of what is known as the reading method. In this case, students' only goal is to become proficient readers in the L2. It was argued that English learners in India, for example, would benefit more from reading than from any other linguistic skill; furthermore, reading was thought to be easier to acquire. The reading method makes use of the first language for explanations, and there is a heavy emphasis on vocabulary development. The reading method also focuses on the use of graded texts, reading strategies, and rapid reading techniques. This method was the first to introduce language instruction for a specific purpose. In some parts of the world, reading in English is still emphasized more than any other skill because university textbooks and many technical and scientific materials are often available only in English.

A major shift took place in the 1950s with the advent of the audiolingual method. This method drew on both the linguistic and psychological theories of the day: structuralism and behaviorism, respectively. Linguists used contrastive analysis, comparing the first language and English, to predict areas of difficulty for language learners; this information was then used to design lessons. A strong reliance on spoken language was a consequence of the linguistic principle that oral language is paramount and that written language is simply a by-product of speech. Behaviorist concepts of habit formation and reinforcement resulted in a heavy emphasis on repetition, mimicry, and memorization of controlled units of language. A typical audiolingual lesson consisted of a dialogue presented to the class by the instructor, using only the L2. The dialogue was repeated by the students several times in chorus, and then in smaller groups. The emphasis was on speaking accurately, both in terms of grammar and pronunciation. Students were given positive feedback for correct productions; errors were avoided by the use of intensive drill. Students were often assigned homework in the language lab where they practiced grammatical patterns. Reading and writing were secondary in importance to the development of good speaking skills. Because of its scientific reputation, as well as heavy financial support from governments and publishers, the audiolingual method became extremely influential.

As a reaction to the behaviorist underpinnings of audiolingualism, the cognitive code approach developed in the late 1960s. The main premise of this approach is the need for the learner to explicitly understand the rules of the language at all levels, rather than simply memorizing dialogues or other chunks of language. Instructors actively teach phonological and grammatical rules, and all four skills (reading, writing, listening, and speaking) are taught together.

Other alternatives to the audiolingual method were classified by H. Douglas Brown as the "designer methods" of the 1970s. These were idiosyncratic and largely atheoretical approaches to language teaching developed by charismatic professionals, most of which are not practiced extensively today. In the Silent Way, developed by Caleb Gattegno, students' exposure to vocabulary is extremely limited in the first month. Words are represented on wall charts, and each letter is color-coded to provide a visual representation of sound and spelling correspondences. Students are encouraged to rely heavily on each other rather than on the teacher, who is to remain as silent as possible. The Silent Way requires considerable training on the part of the teachers.

Another method, developed by Georgi Lozanov, is Suggestopedia, which originated in Bulgaria. In this approach, students are given a new persona. Classes are small and take place in a comfortable setting with easy chairs. The teacher is required to use Baroque music for the séance component of the lesson, during which the students are instructed to use yogic breathing in time with the music while the teacher relays the dialogue in a whisper, a dramatic voice, and in a normal reading. Lozanov claimed that the new personas taken on by the students alleviate a fear of making errors; he also stated that the vocabulary gains with Suggestopedia are far superior to those of any other method.

Like Suggestopedia, one of the aims of community counseling learning, which is based on principles of Rogerian psychology, is to reduce anxiety. The students sit in a circle while the teacher (counselor) stands behind them. When one student chooses to say something to another, he or she utters a sentence in the first language and the counselor translates it into the L2. The student then repeats the translated utterance, and the counselor moves to the recipient to translate a reply. The students then work together on a tape recording of the lesson to work out the linguistic features of the new language.

James Asher, the proponent of the total physical response (TPR) approach to language learning, maintained that physical actions reinforce the learning of vocabulary. In response to a series of the teacher's commands, students are required to act out a sequence of events (much like Gouin's series method). For instance, students may be shown a sequence of actions that mimes washing the dishes. Their instructor then encourages them to respond to commands such as "Put in the plug," "Turn on the tap," "Put soap in the water," "Turn off the tap," and so on. Asher argued that, by acting out the behaviors, students developed a physiological memory of their new language. Asher believed that the acquisition of a L2 was the same as the acquisition of the first language, and he encouraged a silent period in adults that parallels that observed in many children. Once learners feel comfortable with the commands, he claimed, they will start to speak on their own.

ESL Teaching Today

In the early 1980s, communicative language teaching (CLT) became the most popular approach to ESL in North America. At the same time, in Europe the notional-functional approach was widely implemented. Both are still widely practiced today, either in their original form or in an adaptation. The notional-functional approach broke away from a reliance on grammatical structures as the organizing principle for L2 courses and focused instead on functional uses of language, such as greetings, invitations, apologies, opinions, and notions (contexts) such as shopping for groceries.

CLT emerged out of the communicative competence framework established by Canale and Swain and Stephen Krashen's model of L2 acquisition. Krashen argued that L2 acquisition depended on exposing the learner to comprehensible input—that is, language that the learner could understand with the use of contextual clues. He suggested that the input should be at a level just slightly higher than the learner's own proficiency. For this reason, he advocated the use of TPR in the early stages of acquisition, because the meaning of the actions and accompanying language are readily apparent. Perhaps the most contentious aspect of Krashen's theory was the distinction he made between learning and acquisition. Krashen maintained that learning involved consciously putting into memory vocabulary and grammar, whereas acquisition was the subconscious "picking up" of an L2. He claimed not only that learning played a very small part in the overall development of an L2 but also that learning could not turn into acquisition. In other words, learning and acquisition were viewed as two unconnected processes, and learning simply acted as an editor of acquired spoken and written forms.

Another aspect of Krashen's theory that was widely challenged was his claim that the only requirement for mastering an L2 is that the learner is exposed to enough comprehensible input which gradually becomes increasingly complex. Michael Long challenged Krashen's claim, arguing that comprehensible input is a necessary but insufficient component of language acquisition and that interaction, which involves negotiation of meaning, is essential.

Many ESL teachers embraced CLT and designed lessons that provided comprehensible input but did not explicitly teach grammar. Borrowing from notional-functional approaches in Britain, many classroom curricula were developed around themes such as "going to the doctor" and functions such as "making an argument." Often lessons centered on vocabulary development, role playing, and group work.

Although students in CLT classrooms often became relatively fluent in English, they persisted in making grammatical errors. Eventually, applied linguists such as Swain, Richard Schmidt, and Long argued that L2 learners needed a focus on form within a communicative framework. Swain put forward the output hypothesis; she maintained that unless pushed by interlocutors, students analyze their L2 at a semantic level but pay little attention to syntactic structures. By focusing on their own output, that is, their own productions in the L2, they are pushed to become more accurate syntactically because of the feedback they receive. Similarly, Schmidt proposed the noticing hypothesis. He suggested that unless learners have certain linguistic elements brought to their attention, they will not notice them in the input; in other words, a grammatical structure that is not salient or transparent will not be acquired unless teachers instruct the students, either explicitly or by focusing their attention on that form in a task that requires its use.

Recently, some educational theorists have shifted away from viewing L2 learning as a purely psychological phenomenon and toward viewing it as a social construction. The emphasis is on language as a social practice that takes place among people as they negotiate and co-construct meanings in a range of contexts. In pedagogical situations, a sociocultural approach does not focus on language learning as an independent activity. Students are encouraged to interact, using each other as supports while creating collective and collaborative discourse relevant to their own experiences, interests, feelings, and ideas. There is also recognition that identity is inseparable from situational and cultural contexts and language. Thus, social positions and the relative power that they entail are considered to play a key role in language learning processes.

Most of the preceding discussion has focused on language classrooms, but ESL is also taught through content in K–12 settings to immigrant or aboriginal children. There is a variety of delivery methods for school-age children, including pull-out for one-on-one instruction or small group work, separate ESL classrooms, and adapted lessons of subject matter that are appropriate for both ESL and native-speaking students. Bernard Mohan and his colleagues proposed

the knowledge framework approach to adjusting the basic school curriculum to make it accessible to ESL students in mainstream classrooms, while at the same time helping them to recognize how English works. Teachers are encouraged to analyze the subject matter to be taught and the students' texts to determine which knowledge structures (classification, principles, evaluation, description, sequencing, and choice) are evident and which linguistic structures are expressed there. For example, in a science text dealing with a topic such as pollution, there are likely to be discussions of cause and effect (principles) and types of pollution (classification). Each of these knowledge structures is characterized by particular linguistic forms. Cause and effect is often represented by phrases such as *is due to, because of*, and *as a result of.* Mohan advocates the use of key visuals (e.g., time lines, flow charts, tables, diagrams) to help students see the relationships among concepts and the words that express them. Schools that have used this model have found it to be quite successful, but it is extremely time and effort intensive for individual teachers to modify their lessons along these lines.

Two other approaches that are sometimes used to teach English to K–12 learners involve dual language and bilingual programs. The former can be characterized as two-way immersion. In New Mexico, for example, there are several schools in which half the students are native English speakers and half are native Spanish speakers. The students study in both languages and develop oral and literacy skills in both. Bilingual schools are patterned after immersion schools in that subject matter is taught in the students' first language and in English, usually a half day of each language. In Edmonton, Alberta (Canada), children can choose from more than a dozen languages that are taught in conjunction with English.

There is growing concern in some communities that students in K–12 classrooms are not becoming proficient in English. One of the most vocal advocates for ESL students is Jim Cummins, who proposed that children's oral fluency can be misleading because they quickly learn basic interpersonal communication skills, but they take much longer to acquire cognitive academic linguistic proficiency. To achieve academically, students require up to 8 years of language support, but most school districts do not have the resources to provide the programs that ESL students need.

EFL

Most of the methods and approaches described previously in this entry can be found in practice somewhere in the world today in EFL settings. Because the goals of learners in these contexts are so diverse, communicative competence is not necessarily the ultimate objective. Thus, reading courses may be all that is required of students in some situations, because their content textbooks are written in English. However, there is a growing desire for oral proficiency, in some locations, that has resulted in some novel pedagogical developments. In both Korea and Spain, for example, whole English-speaking towns have been established, where native speakers of English go on a paid holiday for the express purpose of talking with Korean or Spanish EFL learners in English. Every encounter is an opportunity for a learner to have an authentic conversation with an English speaker.

Factors That Affect Ultimate Attainment of English

There are several variables that determine whether a learner can become completely proficient in English. Age of learning has long been proposed as having a major influence on ultimate attainment. It was thought that puberty was a critical period for L2 learners; however, research by James Flege and his colleagues showed that there are many individuals who started learning English much earlier than adolescence who are easily identified as L2 speakers. They have proposed that there is a sensitive period very early in life for complete language acquisition. Others have suggested that some adult learners have attained full proficiency; they argue that other factors, such as amount of exposure, motivation, quality of instruction, issues of cultural identity, and native language, have more impact on ultimate attainment than age.

Current Issues

There are several controversial issues in English teaching today, one of which has to do with the variety of English that should be taught, particularly in EFL settings. Some people believe that a standard version of English such as Received Pronunciation (the Queen's English) or standard American English is the best model for communication. Others have argued that the local variety of English should be taught, such as

Indian English in India. The students' purposes for learning English and who they intend to interact with should determine the model that they require.

Another ongoing issue has to do with ESL assessment, which is not only a thriving business but also a high-stakes enterprise for L2 learners. Organizations such as Educational Testing Services produce tests such as the TOEFL (Test of English as a Foreign Language) and the TOEIC (Test of English for International Communication). The TOEFL is used by most postsecondary institutions in North America as an admission criterion for international students. The TOEIC is often required for job promotions in non-English-speaking countries. Because the stakes are so important for test-takers, what they choose to study and what is taught are often dictated by the nature of the tests. In the past, these tests and others focused on discrete elements of English, but they now tend to integrate linguistic skills.

Another matter that concerns many instructors is the question of whether an English teacher should be a native speaker. Nonnative instructors have argued that they have a better understanding of the difficulties faced by their students; they also offer a more realistic model of what can be accomplished. However, in many EFL contexts, employers give preference to native speakers of English, even though many have little or no training. Others argue that native speaker status is immaterial and that there are two fundamental requirements of any good ESL or EFL teacher: high proficiency and high linguistic awareness in English and strong pedagogical skills.

In recent years, there has been a growing recognition that ESL and EFL instructors should be professionals and should meet standards established by national or international organizations. Furthermore, standards have also been developed for programs. There are still many contexts in which English is taught where neither the instructor nor the program has been approved or meets any set of standards, but this will likely change in the coming decades.

Tracey M. Derwing and Judy Cameron

See also Bilingual Education; Bilingualism; Multicultural Classrooms

Further Readings

Brown, H. D. (2007). *Principles of language learning and teaching.* White Plains, NY: Pearson Longman.

Crystal, D. (2003). *English as a global language* (2nd ed.). Cambridge, UK: Cambridge University Press.

Hawkins, M. (2004). *Language learning and teacher education: A sociocultural approach.* Clevedon, UK: Multilingual Matters.

Lightbown, P. M., & Spada, N. (2006). *How languages are learned* (3rd ed.). Oxford, UK: Oxford University Press.

Spada, N. (1997). Form-focussed instruction and second language acquisition: A review of classroom and laboratory research. *Language Teaching*, *30*, 73–87.

Episodic Memory

Episodic memory refers to a division of declarative memory responsible for storing and retrieving information about one's own personal experiences. Episodic memory is the theoretical component of the memory system that allows one to revisit and relive events that occurred within one's personal history. Once experienced, events such as one's first day of school, graduation day, and one's first kiss can each be called into consciousness and relived at a later time because of episodic memory.

Episodic memory is defined in a manner similar to another common memory system, autobiographical memory. Indeed, episodic memory and autobiographical memory are similar enough that some theorists have suggested that the two terms be treated as synonymous. However, this proposal has been met by some resistance from advocates of autobiographical memory, who point out that memory theorists who study autobiographical memory tend to focus their attention on how participants remember events of personal significance, while those interested in episodic memory display far less emphasis on memories for events of autobiographical significance. Some theorists have taken steps to further distinguish between autobiographical memory and episodic memory by listing measurable differences between the two systems. For example, one such difference would be that episodic memories are typically retained for a short period of time, whereas autobiographical memories are retained for a longer time.

Episodic Memory in Educational Psychology

Specific attempts by educational psychologists to study episodic memory have been relatively rare. This

is surprising in that there are some intuitive applications of episodic memory research that those in the education field may find potentially valuable. Although it seems plausible that more advanced knowledge of the application of episodic memory research to education domains might allow for the development of educational curricula that students are more likely to remember, little research has been performed on this idea to date. Recently, some psychologists have argued that an increased emphasis on episodic memory in educational psychology could lead to curricula specifically designed to make it more likely that students remember teaching and learning experiences. Exploration of the intersection between these two active research areas may prove to be a productive avenue for future research.

History of the Concept

In 1972, Endel Tulving proposed a distinction between episodic memory and semantic memory (the memory system responsible for storing facts and general knowledge regarding the way the world works). In his framework, however, episodic memory is not truly independent of semantic memory but rather is a subsystem that is highly dependent on semantic memory.

To help operationalize the differences between episodic memory and other forms of memory, Tulving proposed that we have two ways of accessing the events from our past: remembering and knowing. In Tulving's conceptualization, we "remember" a piece of information from the past when the information is accompanied by the conscious recall of specific details associated with original experience of the information. In contrast, we sometimes remember information from the past but have no conscious recollection of any details about how we learned the information. Tulving referred to this type of access to the past as "knowing."

Upon its introduction, the distinction between episodic memory and semantic memory was readily accepted as a useful heuristic for classifying memory tasks and procedures. However, controversy soon bloomed when it was proposed that episodic memory and semantic memory were functionally and biologically distinct systems.

Evidence for Separate Systems

Tulving and his colleagues have developed a number of lines of evidence intended to differentiate episodic memory from semantic memory. These lines of evidence range from the phenomenological, to the traditionally behavioral, to the physiological. On the phenomenological end of the evidence, Tulving points out that episodic memories have an affective tone associated with them that is missing in semantic memory. A more traditionally behavioral area where episodic memory has been theorized to dissociate from semantic memory is in how likely subjects are to lose information from the respective systems. Some have proposed that episodic memory is more vulnerable to forgetting than is semantic memory.

Physiological evidence for a distinction between episodic and semantic memory systems comes from neuropsychological case studies. The best example of these case studies comes from patient K.C. (amnesic case studies are frequently referred to by their initials in the neuropsychological literature to provide a degree of protection for their anonymity). K.C. exhibits symptoms consistent with a distinction between episodic memory and semantic memory in that his episodic memory was almost entirely destroyed in a motorcycle accident, while his semantic memory remains relatively intact. K.C cannot recall any personal experiences but is able to learn new facts. Although a handful of similar cases exist, brain damage similar to that observed in K.C. is extraordinarily rare. These cases demonstrate that it is possible to have a functional semantic memory in the absence of episodic memory.

Functional neuroimaging studies have been used to identify the physiological underpinnings of episodic memory. Distinct patterns of cortical activation associated specifically with episodic memory have been identified. Neuroimaging studies have typically found medial and left lateralized activation patterns associated with retrieval of episodic memories. In particular, the medial frontal cortex and left hippocampus tend to be active during episodic memory retrieval.

Criticisms of the Separate Systems View

Although the previously discussed evidence has been interpreted by Tulving and others as support for the idea that episodic memory is a memory system distinct from semantic memory, some have argued that this distinction is unnecessary. In keeping with the scientific value of parsimony, it has been argued that there may not be a need to hypothesize multiple

memory systems. Indeed, several of the studies cited as evidence for a distinction between semantic memory and episodic memory have been questioned by the theory's critics. For example, critics note that the hypothesized distinction between forgetting rates in episodic memory and semantic memory does not seem to be supported by the data.

Research

Since the episodic memory distinction was first proposed in 1972, hundreds of studies of episodic memory have been conducted, and episodic memory remains an active area of research today. Although memory theorists continue to debate whether episodic memory is a truly separate system from semantic memory, most agree that the concept of episodic memory has utility.

Jeremy K. Miller

See also Long-Term Memory; Memory

Further Readings

Maguire, E. A. (2002). Neuroimaging studies of autobiographical event memory. In A. Baddeley, M. Conway, & J. Aggleton (Eds.), *Episodic memory: New directions in research* (pp. 164–180). New York: Oxford University Press.

Martin, J. (1993). Episodic memory: A neglected phenomenon in the psychology of education. *Educational Psychologist*, *28*, 169–183.

McKoon, G., Ratcliff, R., & Dell, G. S. (1986). A criticism of the semantic-episodic distinction. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*, 295–306.

Tulving, E. (1972). Epsiodic and semantic memory. In E Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 381–403). New York: Academic Press.

Tulving, E. (1983). *Elements of episodic memory.* New York: Oxford University Press.

Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology*, *53*, 1–25.

EQUILIBRATION

Developmental theorists make a distinction between developmental states and developmental mechanisms. *Developmental states* are chronologically arranged snapshots of children's skill levels at particular points in time. An example of a sequence of developmental states would be the four stages of cognitive development that were proposed by Jean Piaget in his classic theory (i.e., sensori-motor, preoperations, concrete operations, formal operations). As children enter a new stage, they acquire a new set of skills and move to a new level of performance. *Developmental mechanisms*, in contrast, are the factors that are used to explain children's progression through developmental states (i.e., why a child progresses from one state to another in a sequence). The primary developmental mechanism in Piaget's theory is called equilibration. *Equilibration* pertains to restoring the balance between two competing tendencies in the mind: assimilation and accommodation. Piaget used the notion of *assimilation* to describe the process of incorporating experiences and information into existing knowledge structures in the mind. Metaphorically, children find a "home" for this information in their existing knowledge structures. To say that a child has assimilated an idea or experience is to say that he or she understood the idea or experience. Piaget thought that mental assimilation was analogous to the biological assimilation that takes place when the human body extracts what it needs from food and incorporates the extracted nutrients into existing organs and tissues.

Sometimes an idea is so discrepant from what a child believes or knows that it cannot be assimilated. Piaget used the notion of *accommodation* to describe the process of changing the existing configuration of knowledge in the mind in order that the troublesome idea can be assimilated. In most cases, assimilation is always partial in the sense that children only assimilate that portion of an experience that is consistent with their current understanding. And unless an experience or idea is identical to previous ones, every act of assimilation usually precipitates accommodation of knowledge as well. For example, when a child encounters a new species of dog for the first time and is told that it is a dog, this information finds a home in the existing network of ideas (i.e., it is assimilated), but the network is also changed as the child's mind creates a new representation corresponding to the new subtype of dog (i.e., there is accommodation).

In his book, *Play, Dreams, and Imitation*, Piaget argued that "imitation is a continuation of accommodation, play is a continuation of assimilation, and intelligence is a harmonious combination of the two" (p. 104). This quote suggests that when one assimilates, one inserts one's own ideas into reality; when

one accommodates, one's schemes and ideas come into closer conformity with reality and tend to be fairly direct copies of it. Play and fantasy, moreover, are examples of overassimilation (i.e., putting too much of one's ideas into reality). In contrast, children engage in overaccommodation when they try to directly copy the actions of someone (without putting their own "spin" on the actions).

To illustrate assimilation and accommodation further, consider the following example. Young preschoolers who are passengers in their parents' cars often think that the moon is following them when they drive at night. The physics of the explanation of what is actually happening is too abstract for young children to comprehend, so they could not assimilate even this explanation if it were provided. Ultimately, however, their knowledge of the physical world will change enough that they may eventually understand the explanation as young adults.

Piaget argued that confronting discrepant ideas is absolutely essential for knowledge growth. If children never had experiences or heard information that contradicted the erroneous ideas that they construct by themselves, they would never develop the correct conceptions. Thus, the "readiness" idea of waiting until a child's mind matures enough to teach a topic to them is actually an implication of nativism, not Piaget's constructivism. Piaget did not believe that ideas were inborn as nativists do; rather, he believed that children build up their ideas step by step.

Moreover, Piaget argued that assimilation and accommodation work in opposition to each other. The central tendency for assimilation is to keep the existing knowledge structure the same and find a place for new information in this structure. The central tendency for accommodation, in contrast, is to change the existing knowledge structure. It is not possible to keep things the same and change them at the same time. Thus, only one or the other—assimilation or accommodation—"wins out" in a given situation. This "battle" between these processes means that change in children's misconceptions can be frustratingly slow for teachers. That is, even when a misconception is repeatedly pointed out and explained to children, they may cling to the erroneous idea for some time. Perhaps this implies that assimilation normally takes precedence over accommodation. When the battle is resolved over some idea and a balance is restored between assimilation and accommodation (i.e., equilibration has occurred), children's

understanding usually moves to a higher plane and often becomes more abstract as well. For example, for children to come to understand that dogs and people are both animals, they have to change their concept of "animal" in such a way that it is more abstract (e.g., "a living thing that can move itself" from "furry four-legged things").

Piaget argued that equilibration was particularly adaptive because the changes that are incorporated into mental structures are such that the mental system of ideas becomes capable of anticipating future problems that might arise. That is, the change not only allows a child to assimilate a particular idea that once was difficult to grasp but also helps the child grasp related ideas that he or she has not yet encountered. In this way, equilibration works in a manner analogous to how the immune system creates structures (i.e., antibodies) that can deal with problems that arise in the future (i.e., reoccurrence of a virus that was encountered before).

James P. Byrnes

See also Cognitive Development and School Readiness; Constructivism; Piaget's Theory of Cognitive Development

Further Readings

Chapman, M. (1988). *Constructive evolution. Origins and development of Piaget's thought.* Cambridge, UK: Cambridge University Press.

Gruber, H. E., & Vonèche, J. (1995). *The essential Piaget.* Northvale, NJ: Aronson.

Piaget, J. (1952). *The origins of intelligence in children.* New York: International Universities Press.

Piaget, J. (1962). *Play, dreams, and imitation in childhood.* New York: W. W. Norton.

Piaget, J. (1983). Piaget's theory. In P. Mussen (Series Ed.) & R. Lerner (Vol. Ed.), *Handbook of child psychology* (4th ed., Vol. 1). New York: Wiley.

Piaget, J., & Inhelder, B. (1969). *The psychology of the child.* New York: Basic Books.

ERIKSON'S THEORY OF PSYCHOSOCIAL DEVELOPMENT

In one biography, Erik Erikson is referred to as an architect of identity. He was trained in psychoanalysis, and one of his main scholarly interests was

describing the nature of humankind. He engaged in personal and cultural observations in his theoretical writings about identity formation. His major contribution was the formulation of an elaborate life-span theory of identity formation based on eight life stages. Each stage is based on age-appropriate and culturally defined crises, or turning points, which demand that the individual choose between dialectic and opposing developmental outcomes. Each stage outcome provides a contribution to identity formation in the form of ego strengths.

Erik Erikson's theory of psychosocial development has been useful in helping teachers and educators to understand child development for more than 50 years. It remains a solid theoretical perspective for understanding child and adolescent development and adult learning. There is an ever-expanding body of research that demonstrates Erikson's many ideas can be empirically supported as being valid and practical.

Definitions of Identity

Identity is defined in many ways. Most often Erikson refers to *ego identity*, or the portion of personality that functions to direct, guide, and select thought and action. It is the part of personality that refers to the consistency and sameness that a person uses as a style of individuality and its meaning about the self for significant others like family members, teachers, or friends. Identity can be thought of as a form of relatively persistent character one shares with others.

In addition to ego identity, Erikson also refers to the stable and coherent, but yet evolving, character one develops with a group's ideals in the form of social identity. Social identity is most readily observed in the expectations and commitments to the values, goals, and ideals of institutions such as religion, politics, occupation, and the family. In particular, social identity is based on the ideology that is formed within a commitment to a particular institutional value. For example, an occupational identity based on the achievement motive of capitalism differs from an occupational identity based on the conformist motive of a communism ideology. Religious identities differ within Protestant, Catholic, Islamic, Hindu, Jewish, or other forms of religious social ideologies. Likewise, political identities vary among party, state, or group ideologies.

Families vary as to the values, goals, and aspirations that parents have for their children, and identification with family ideologies will vary widely within any given community. Many factors will contribute to whether youth will identify (or not) with their own family values and ideologies. Indeed, individuals strive toward psychology unity, where various domains of identity across the various institutional ideologies merge into a comfortable coherent and consistent sense of individuality. Extreme differences in values or incompatible goals or values in adopted commitments create angst or anxiety that links with self-consciousness and discomfort, leaving a person with a sense of fragmentation. A fragmented self (or incompatible elements within either an ego identity or a social identity) will lead to uncomfortable states of anxiety and a sense of inner conflict.

Eight Life Stages

The eight life stages that facilitate identity formation are set by societal expectations to accomplish specific developmental tasks. Each of these tasks, as it is successfully resolved, provides an innovative strength to both ego- and social-identity development. These stages are based on an epigenetic principle wherein the individual is driven toward, is aware of, and interacts with life dilemmas. Human growth is based on a dialectic struggle of inner (psychological) and outer (social) forces where each stage of life involves a crisis or turning point. Each of the life stages provides a critical point to increase potential for growth in identity based on an ontogenetic evolution. Healthy development is based on the assumption that every stage comes to its ascendance: A crisis is resolved and is, ideally, concluded with an effective and meaningful identity resolution. To complete a life stage, a person must synthesis (or resynthesize) early childhood identifications for novel and newly constructed identifications. Put another way, the individual must both selectively repudiate and assimilate childhood identifications with new configurations of ego identity.

Each stage of life consists of issues regarding a designated age or stage, questions and virtues, appropriate precursors, contributions to identity, and a radius of social order. Erikson begins with infancy and a crisis of trust versus mistrust. In general terms the question to be answered is "What do I have and what can I give?" The virtue is hope. In early childhood, the crisis is autonomy versus shame and doubt. The question is "What can I will freely?" and the virtue is will. Somewhat later in childhood the crisis becomes

initiative versus guilt, and the question is "What can I imagine I will be?" The virtue is purpose. For the school-age child, the crisis is industry versus inferiority, and the question is "What can I learn to make?" The virtue is competence. In adolescence, the crisis is identity versus role confusion. The question is "Who do I choose to be ideologically and occupationally?" The virtue is fidelity. In young adulthood, the question the crisis is intimacy versus isolation. The question focuses on "What I can give to another?" The virtue is love. In middle adulthood, the crisis is generativity versus stagnation. The question is "What I can give and care for?" The virtue is care. As one approaches old age, the crisis is integrity versus despair. The question is "What is the essence of my existence?" The virtue is wisdom.

The radius of social order begins with the important power of the maternal person and expands through school, community, technology, and beyond. In infancy, the contribution to identity is time perspective versus time diffusion, and the radius is the maternal person. In early childhood, the contribution to identity is self-certainty versus identity consciousness, and the radius is parental care. In childhood, the contribution is role experimentation versus negative identity, and the radius is the basic family. For the school-age child, the contribution is the anticipation of achievement, and the social radius is the neighborhood, school, and technology. In adolescence, the contribution is identity consolidation, and the social radius includes peer groups, models of leadership, and ideological perspectives of the social institutions associated with occupation, family, religion, and politics.

Importance of Psychosocial Moratorium

Each of the first five life stages provides an important precursor to identity formation, and the remaining three offer further consolidation through the virtues of love, care, and wisdom. However, Western societies, in particular, provide an extended period of time for role experimentation to explore identity options and select and make commitments to those identity fragments that will become integrated into a uniformed sense of ego identity and social identity. This period of time is called a *psychosocial moratorium*. It is a period when schools, parents, and communities allow for experimentation and offer opportunities for teenagers and sometimes even emerging adults to explore what they can do to select and prepare for an occupation, profession, or career. During this time, youth can find many teachers, parents, adults, and others ready to role model and assist in preparing them through education, mentoring, and apprenticeship experiences. Failure of youth to effectively utilize the psychosocial moratorium can result in a negative identity whereby youth identify with negative features or characteristics of society that cause the youth to appear rebellious or alienated.

Women's Identity

The most controversial aspect of Erikson's theory of identity development is his writings on gender differences. In early psychoanalytic writings based on observations during play therapy, boys were observed to use play building blocks to make towers, whereas girls made circles or walled enclosures. This led to notions of boys using outer or phallic space and girls using inner biological space in their psychology development and to the belief that girls' development was based, in large part, on their biological destiny of procreation and motherhood. Thus, identity formation for men occurs with identity formation functioning clearly as a precursor to intimacy formation. In contrast, for women, identity formation unfolds concurrently with intimacy formation. There has been ample critique of this controversial statement on gender difference; however, there is little strong empirical research evidence of a longitudinal nature to establish whether or not this hypothesis is accurate. Most of the evidence that has been presented to indicate Erikson's notion of biological destiny for women's identity formation trajectory has been of a qualitative nature and could easily be interpreted arguably as supporting or not supporting this assumption. Nonetheless, the position on biological determinism was maintained by Erikson throughout all of his lectures, writings, and interviews. In a historical period of heightened feminism, this position was unpopular and often criticized.

Identity Statuses

Theoretical perspectives are often difficult to define and measure. Only limited aspects of Erikson's theory of psychosocial development have been defined well enough to measure. The two stages most widely studied have been *identity versus role confusion* and *intimacy versus stagnation*.

James Marcia identified two dimensions from Erikson's writings that are part of the developmental process of identity formation. These dimensions include a searching or crisis process and an outcome of commitment. Using these two dimensions and crossing them at the midpoint in the middle he created four states (which he called *statuses*): no searching and no commitment, no searching and commitment, searching and no commitment, and searching and commitment. In his framework, searching, or *identity crisis*, involved active attempts to formulate or construct an identity. *Commitment* refers to the construction and formulation of a comfortable sense of identity. No searching and no commitment is called *identity diffusion.* No searching but with commitment is called *foreclosure*, a form of identity where there is a passive acceptance of identity of a previous generation or parents. Searching but no reported commitment is called *moratorium*, an individual who is in the process of actively trying to formulate a sense of self through discovery and or self-creation. Finally, evidence of past searching and current commitment is called *identity achievement*. Clinical and social science research reveals that adolescents and youth with diffused ego identity are most likely to have mental health problems, whereas their peers who are identity achieved are most likely to have the best mental health of the four types of identity status types. These findings are fully consistent with Erikson's theory of ego-identity formation.

Jacob Orlofsky and colleagues used three major criteria for assessing a person's level of intimacy as suggested by Erikson:

1. Does the person have close relationships with male and female friends?

2. Does he or she have an enduring heterosexual relationship?

3. Are the person's close relationships deep or superficial? (Depth includes openness, affection, respect, loyalty, a capacity to accept and resolve differences, and mutuality.)

Based on these criteria, five relationships, or intimacy statuses, can be identified through interview techniques. At the two extremes are the intimate and the isolated statuses. The intimacy status involves love and an enduring commitment. In the isolated status, the person is withdrawn and has no close personal associations. Between these two extremes are three other types. The preintimate status is a loving relationship but with no enduring commitment. A stereotyped status is an association with male or female friends that lacks depth and closeness. The pseudointimate status is artificial; it only seems to have depth and caring.

Wide use of the intimacy and identity statuses has revealed that individuals who are identity diffused are also intimacy isolated. In contrast, identity-achieved individuals have considerably more intimate relationships. This is again consistent with Erikson's notions of interstage linkages among identity and intimacy resolutions in the fifth and six stages of ego-identity development.

Power of Educational Institutions

Social institutions have a variety of processes by which they motivate and encourage individuals to develop. The strongest influence is through the ascendance of age-based crisis points where the expectation is that the individual will make a turning point by resolving a challenge between two opposing oppositions (e.g., identity formation versus role confusion) during the teenage and emerging adult years. Work and educational systems provide mentors and teachers to help guide and direct youth to make choices through role modeling, instruction, encouragement, and other positive educational techniques. Erikson indicates that the best teachers are ones that know how to balance play with work to make the decision a pleasurable and positive experience. Educational research clearly demonstrates that positive instructors and peers enhance identity formation among students during the high school and college years.

Gerald R. Adams

See also Identity Development; Psychoanalytic Theory; Psychosocial Development; Social Development

Further Readings

Coleman, R. (Ed.). (2000). *The Erik Erikson reader.* New York: W. W. Norton.

Erikson, E. H. (1963). *Childhood and society.* New York: W.W. Norton.

Friedman, L. J. (1999). *Identity's architect: A biography of Erik H. Erikson.* New York: Scribner.

ESSAY TESTS

Essay tests are a type of constructed-response assessment in which test-takers display skills and knowledge by writing responses to an assigned task, question, or prompt. Instead of selecting answers to multiple-choice questions, the test-taker may be required to present an opinion on an issue, explain a chemical process, describe a scene in a foreign language, present conclusions drawn from a set of economic data, write a narrative, or demonstrate other skills and understandings in writing. The tasks may be simple or highly complex, discrete or multiphased; the response format may vary from a paragraph to a letter, a multiparagraph essay, a proposal, and so on.

For centuries, essay testing was the primary form of assessment. But with an emphasis on statistical rigor and the advent of rapid-scanning equipment and other technical advances in the mid-20th century, selected-response tests (primarily multiple-choice tests) emerged as a generally more objective, efficient, and cost-effective form for large-scale assessments. Selected-response tests usually sample a broader content domain: In an hour of testing time, test-takers might respond to 45 multiple-choice questions instead of only one or a few essay questions. Furthermore, the more questions in a test, the less a test-taker's performance on any one question influences the final score. These factors, in addition to the right-or-wrong nature of the questions, make selected-response tests "more reliable"—that is, test-takers' scores are substantially more consistent across different versions of the same test. As selected-response testing evolved, essay testing was often dismissed as more subjective and less reliable.

In recent decades, however, developments in cognitive psychology and an increased emphasis on more authentic assessment have helped spur interest in having test-takers *display* or *perform* certain skills. Essay testing has become a highly rigorous discipline supported with research, technological advancements, and published guidelines for best practices (although theorists continue to debate conceptual issues, including the extent to which test reliability should be emphasized or readers should set aside their personal reactions to essay responses).

Essay testing requires careful deliberation at every stage. The process of specifying the domain of content and skills to be assessed often includes surveys, focus-group sessions, committee deliberations, or reviews of professional and curriculum standards. The development process may also include prototype testing of various formats, directions, topics, and evaluation criteria, as well as questionnaires or interviews ascertaining test-takers' opinions about the tasks. Below are some important steps in designing an essay test:

- Determine which skills and knowledge are best assessed via essay tasks.
- Consider the extent to which writing fluency could unduly affect measurement of other skills or knowledge.
- Include a sufficient number of test questions—possibly combining essay tasks and selected-response questions—for desired content coverage and reliable scores.
- Provide adequate time for each task.
- Consider how requiring test-takers either to hand-write or word process essay responses may affect test-takers' ability to display the skills being assessed.
- Give test-takers ample opportunity to prepare by prepublishing information about the task types and scoring standards, including sample questions and sample responses with scores and commentary.
- Determine the extent to which evidence supports the claim that the essay test is valid for its intended purpose—that is, that the inferences and actions based on test scores will be appropriate, meaningful, and useful. For example, conduct studies to investigate relationships between test-takers' performance on the essay test and demonstrations of the same skills in other contexts.

In essay testing, how responses are scored becomes an integral part of test design. First, the scoring methodology determines what type of information will be obtained. One common approach is *holistic* scoring, in which a reader appraises the overall quality of a response to assign a single score. Holistic scoring can reflect the integrated nature of proficiencies such as writing, reading comprehension, and/or critical-thinking skills; moreover, it is an efficient approach to scoring large volumes of responses. However, it does not provide the diagnostic feedback or detailed information that may be gained from more time-consuming *analytic* scoring, in which a reader may evaluate the different features or traits of a response separately or else award scores based on the presence or absence of specified components (e.g., particular facts about a chemical process). In practice, testing programs may employ variants of holistic or analytic scoring.

Second, to make scoring more reliable, it is important to develop explicit and carefully crafted evaluation guidelines: the *rubrics* or scoring guides that usually require the reader to apply a variety of specifically defined criteria (e.g., accuracy of information, aptness of examples, clarity of expression). Ideally, these criteria embody the consensus of the relevant community of educators and readers. Rubrics also stipulate the scale of numeric or descriptive score levels that readers can assign.

For large-scale testing programs, operational scoring procedures are usually highly structured and standardized. Scoring sessions may be *centralized*, with all the readers gathered together, or *distributed*, with readers geographically dispersed. Distributed scoring can be done online, with readers accessing training materials and scoring responses via an Internet application. Online systems bring great efficiency to the scoring process and allow the scoring leadership to monitor it using up-to-the-minute data on reader performance, including score distributions and rates of interreader agreement.

Scoring protocols that help promote fairness and accuracy include training of readers both before and throughout the scoring session as well as mentoring and monitoring by experienced readers who serve as scoring leaders. Training should familiarize readers with the principles of the scoring methodology, the rubrics, the essay tasks, and the exemplar responses that, for each topic or question, help define each score level in the rubric. As part of the protocols, many testing programs enhance score reliability by requiring that two readers independently score each response, with a third independent reading to resolve cases in which, according to program standards, the initial two scores differ too widely.

The field of essay testing continues to evolve. Recent technological developments include automated scoring by software applications that use semantic analysis or other types of computational approaches intended to emulate the judgments of human readers. Ongoing research and refinements in essay testing are also making it possible to assess a wider range of skills and knowledge and enhance the authenticity of essay tasks.

Mary E. Fowles and Nora V. Odendahl

See also Assessment; Multiple-Choice Tests; Reliability; Validity

Further Readings

Breland, H. M., Bridgeman, B., & Fowles, M. E. (1999). *Writing assessment in admission to higher education: Review and framework* (ETS Rep. No. RR-99-3). Princeton, NJ: Educational Testing Service.

Educational Testing Service. (2005). *ETS guidelines for constructed-response and other performance assessments* (ETS Publication No. 727534). Princeton, NJ: Author.

Martinez, M. E. (1999). Cognition and the question of test item format. *Educational Psychologist 34*, 207–218.

Walstad, W. B., & Watts, M. (2006). Testing for depth of understanding in economics using essay questions. *Journal of Economic Education*, *37*(1), 38–47.

Wolcott, W. (with Legg, S. M.). (1998). *An overview of writing assessment: Theory, research, and practice.* Urbana, IL: National Council of Teachers of English.

ETHICS AND RESEARCH

The knowledge base in educational psychology has been developed by individuals who have conducted research to uncover truths about the affective, cognitive, and behavioral attributes of learners. The growth of any discipline is dependent upon the worthiness of the research that it produces. Research is defined as systematic investigation, including research development, testing, and evaluation, designed to develop or contribute to generalizable knowledge. Effective research must meet the criteria of being relevant and ethically produced. When research does not live up to the standards of defined practice, research misconduct inquiries and investigations are used by institutions to identify the objectionable practice and to remediate the affects.

Research ethics is a broad term that is used to describe adherence to the shared values of efficiency, truthfulness, accuracy, and objectivity. Ethical research must use resources wisely and convey information truthfully by carefully reporting results in a truthful manner without bias. A more narrowly focused term—*responsible conduct of research*—relates to more definable practices that prescribe the research behavior required of investigators.

Responsible conduct of research must take into consideration the following eight elements:

1. Data acquisition and management

2. Mentor–trainee relationships

3. Publication practices and authorship

4. Peer review practices

5. Collaborative research activities

6. Protection of human subjects

7. Research involving animals

8. Management of conflict of interest and commitment

More specifically, responsible conduct of research refers to meeting the requirements relating to those concepts mandated by Congress and the regulations promulgated by federal agencies. However, the broader view of research ethics goes beyond what is required and relates to what should be done regarding moral dilemmas that surround how to make the best research decision.

The federal regulations relating to responsible conduct of research stem from three acts passed by Congress: the 1966 Animal Welfare Act (Public Law [P.L.] 89–544), the 1974 National Research Act (P.L. 93–348), and the 1985 Health Research Extension Act (P.L. 99–158). The 1985 Health Research Extension Act resulted, in 1989, in the establishment of the Office of Scientific Integrity (OSI) and the Office of Scientific Integrity Review (OSIR). In 1992 the Office of Research Integrity (ORI) was established and took over the activities of the former two offices, including the monitoring of research misconduct.

In addition to the regulations developed from these acts, the executive branch agencies of the government have developed guidelines and policies that represent requirements relating to their normal operation. For example, the National Institutes of Health has issued policies for grant awards, and the ORI has issued a Model Policy and Procedures for Responding to Allegations of Scientific Misconduct.

Beyond the federal policies are disciplinary codes of ethics and institutional policies. Most disciplines have outlined principles and, in some cases, appropriate practices that relate to the ethical practice of research. Universities and other institutions must also have research policies if they are to receive federal funds. Again, these policies may go beyond the recommendations of the federal guidelines. Many of these policies amplify and implement the federal guidelines that make up the various elements of responsible conduct of research.

Data Acquisition and Management

Issues relating to data acquisition and management include collection of information, storage, ownership of data, and sharing of data.

Collection of information is governed by principles of effective research. The quality of research depends upon the use of appropriate research methods. Investigators need to be precise in the collection and handling of data. The results must be carefully recorded, analyzed, and interpreted. Data can be recorded in numbered bound notebooks, by permanently dated electronic files, or other permanent means.

Information should be stored in a safe place that is not subject to the elements or catastrophes. Electronic files should be backed up and stored away from the original data. Other materials, such as samples or videos, should be protected so that they do not degrade. Confidential information must be stored in areas that prevent unintended access.

Ownership of data that is generated by researchers is a complex matter. There are various parties that may have claims on the ownership of data. The university or the institution that employs the researcher and governmental or foundation funding sources have a stake in the data collected. Researchers must report the results of their work to the granting agencies, but the control of the data generated is the responsibility of the institution that received the funding. If the funding source was a contract, the researcher will be obligated to provide a product at the end of the project. The organization that received the funding has responsibilities to maintain the data collected. Researchers who leave employment at one institution and take a position at another must seek approval to relocate the data that they have collected.

Mentor–Trainee Relationships

Responsible conduct of research includes developing mentor–trainee relationships. Successful researchers have an obligation to foster the development of others who are learning to become researchers. It is the responsibility of a good mentor to help the student capitalize on the educational experience and to acculturate the student into the practices and ethics of the profession. A mentorship relationship requires that the mentor provide a supportive learning environment. The mentor must devote a sufficient amount of time to facilitate the development of the trainee. This will include providing training, supervision, and honest appraisal. The mentor, through example, will teach the basics of data collection, interpretation, and publication and will give due credit to the trainee for his or her contributions. The mentor will help the student

understand the values of the profession and how to maneuver within the accepted guidelines. Included in the culture of the profession is the way it treats authorship and publication practices.

Publication Practices and Authorship

Modern copyright first appeared in Britain in the 17th century to protect the printing industry from others who would duplicate their products. According to Corynne McSherry, eventually the concept of copyright came to encompass protection of the intellectual property of the author.

At most universities, faculty are assigned the copyright for their works of scholarship unless those rights relate to a patentable product. This policy allows faculty to publish freely without restriction and disseminate to the public the results of their research and scholarship. In situations in which research leads to a new discovery that can be patented, publication may be restricted until a patent can be assigned. A patent prevents others from practicing the invention and gives the developer the right to practice the invention or seek licensing allowing others to commercialize the discovery. Premature publication may prevent the establishment of a patent.

Early disclosure of results of research is often provided at professional meetings and is later further documented in scholarly publications. Fair practices of publication require that all documents provide a detailed and objective description of the research that was conducted. It should describe the methods used, indicate the results discovered, and provide an open assessment of the findings.

When individuals collaborate on a publication, it is often unclear as to who will be credited as the main author and who will have a lesser role. Usually authors are listed in publications in the order of the importance of their contributions. However, whereas in some disciplines the most important contributor is listed first, in others this individual is listed last. Disputes can arise regarding the order of authorship when the authors have failed to discuss this issue beforehand. In psychology the code of ethics states psychologists must take responsibility and credit, including authorship credit, only for work they have actually performed or to which they have substantially contributed. In addition, principal authorship and other publication credits must accurately reflect the relative scientific or professional contributions of the individuals

involved, regardless of their relative status. The code goes on to say that minor contributions to the research or to the writing for publications are acknowledged appropriately, such as in footnotes or in an introductory statement.

Although it is recognized that only individuals who have made significant intellectual contributions to an article should be listed as authors, there is no universal understanding as to what constitutes a significant contribution. In an attempt to bring clarity to this ambiguity, the International Committee of Medical Journal Editors (ICMJE) proposed a set of guidelines that are often used. The ICMJE declared that authors must have taken responsibility for portions of the article's content. This could include contributions to the development and design, acquisition of data, or analysis and interpretation of data. The author must have had a role in the drafting of the article or be involved in the critical revision of the content. The author must also be in a position to approve the final version to be published. According to the ICMJE, acquisition of funding, collection of data, or general supervision of the research group would by itself qualify an individual as an author without meeting the previously stated criteria. The ICMJE further stated that the order of authorship should be a joint decision of the coauthors. The greatest contribution of the ICMJE was to suggest that authors should provide a description of what each person has contributed to the article and that others who did not meet these criteria for authorship but had contributed to the article should be acknowledged but not listed as authors.

Responsible publication practices dictate that the same research results should not be repeated in different publications without citing the previous publication. Duplicate publications have the result of wasting time and resources and also tend to overemphasize results beyond their actual importance. A review of the literature would unfairly weigh the results of the research in favor of the author's conclusions when it has been repeated in multiple journals.

Peer Review Practices

Once an article has been written, it is customary to submit it to a journal for peer review by colleagues who have expertise in the topic and can provide a fair review of the manuscript. Peer review helps determine which research is reliable, valid, and should be published and which research has flaws and should

not be made available to the profession. It is a universal process that takes place in all countries and has the effect of professionals monitoring their own professions. In the same way, peer review is also a part of the process that will determine which research will be funded. Grant reviewers will evaluate the concept, methods of research, budget, and potential outcomes of the research.

Peer review becomes an essential ingredient in the career of the individual and influences the direction of research an individual takes within the profession of educational psychology. With so much at stake, it is therefore important that the process be as open as possible and that it have certain safeguards. Peer review must be conducted in a manner that respects the confidentiality of the research and is free from the reviewer's personal bias. The feedback must be critical yet constructive in its nature so that the author may benefit and so that the profession can be protected. It is also necessary that it be completed within deadlines that permit publication in a timely fashion.

Scholars within a profession have the obligation of investing time to review grants and publications within the profession. Individual reviewers should know their limitations when reviewing. They should review only projects for which they are adequately qualified to render an informed opinion. In addition, reviewers should be familiar with the literature in the field, be able to judge the quality of the manuscript, understand the methodology implemented, be knowledgeable about the type of analysis employed, and be capable of evaluating whether or not the conclusion is supported by the evidence. Beyond these requirements, the reviewer must be able to identify any possible conflicts of interests that might exist. If a reviewer is working on the same research as the author, he or she must disclose this conflict to the editors and either excuse him- or herself or be certain to review the material objectively and not use information learned from the review in his or her own research.

Collaborative Research

Collaborative research has many rewards but poses interesting challenges. Before research begins, collaborators must have the same understanding of the requirements, and thus, a uniform level of training and supervision must be in place. This may include an understanding of welfare of human subjects, animal subjects, and use of hazardous materials.

When researchers are working together with others, either on a publication or a grant, the roles and responsibilities of each researcher must be clearly delineated. Collaborative activities require the clarification of the following issues: Who will submit the funding proposal? Who will submit the proposal protocol for evaluation to a subject review board? Who will store the collected data? Who will have the authority to make changes in the project? Who will write the first draft of an article, and who will report the results of the article? How will intellectual property rights be distributed? These issues can be resolved only if effective communication is established among the collaborators.

If the collaborative research entails external funding, financial management can also become an issue. The Office of Management and Budget produces two documents, Circular A-21 and A-110, which provide guidance regarding the financial management. All members of the collaborative team need to be familiar with these documents so that they are aware of allowable expenditures, equipment use, and reporting procedures.

When research involves the use of materials, it is necessary to formulate agreements that address how such materials will be transferred between researchers and their institutions. This is usually accomplished through transfer agreements that stipulate the conditions under which the materials can be used and information that can be disclosed as a result of their usage.

Protection of Human Subjects

Research that involves human subjects must be designed and conducted in a way that maximizes benefits to the subjects while minimizing risks. The process of ensuring ethical treatment of subjects requires that all proposed research be scrutinized by a human subjects review board.

Research practices in the past that failed to protect subjects led to the need for a formalized method for evaluating and approving human subject research. Much of the effort toward protecting human subjects occurred following World War II. Evidence of atrocities committed during the war and attributed to research activity led to regulations known as the Nuremberg Code. This code and the later Declaration of Helsinki established an international framework that defines ethical treatment of human subjects. The Nuremberg Code established basic principles that make

up the expectations for ethical treatment of human subjects. They include the requirement of obtaining voluntary consent of the subject, establishing the benefit of the study in relation to risk, avoiding unnecessary suffering and injury, requiring adequate facilities and qualified experimenters, and allowing the subjects to discontinue involvement at any time.

Even though such guidelines existed as early as 1947, unconscionable research nonetheless occurred following the war. In the United States, research on the effects of radiation was conducted on subjects in hospitals without their consent. In the 1950s and 1960s experiments on the effects of untreated syphilis were conducted on unsuspecting African American subjects by the U.S. Public Health Service in Tuskegee, Alabama. The resulting concern over these and other violations of subjects' rights led to governmental policy. In 1974 three sets of guidelines were developed by the U.S. government. First, the Congress demanded that the Department of Health, Education and Welfare (now Department of Health and Human Services) further address standards for involving human subjects in research. The resulting outcome was Title 45 of the Code of Federal Regulations, Part 46 (45 CFR 46). In a similar time frame, the Food and Drug Administration established its own rules for human subject research (21 CFR 50 and 56). Also in 1974 the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research was established. Following these efforts, in 1991 many federal governmental departments and agencies adopted a uniform set of regulations known as the Common Rule, which makes up the guidelines affecting human subject research (45 CFR 46, Subpart A-D). The Common Rule requires that studies involving human subjects be approved by a human subject review board or other appropriate institutional officials. The review board also should provide a continuing review, at least once per year, of the approved research. In some situations, the human subjects review board may identify research as being exempt from the federal requirements. The responsibility of enforcing the Health and Human Services regulations falls upon the Office for Human Research Protections in the Office of Public Health and Sciences. To help ensure that researchers understand the principles of protection of human subjects, the National Institutes of Health requires that investigators who submit a grant application receive training on the protection of subjects. The Common Rule requires that an institution that is engaged in federally funded human subject research must file a Federal-Wide Assurance of Protection for Human Subjects.

Decisions regarding ethical treatment most directly spring from the 1979 Belmont Report of the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. The principles identified in the report include (a) respect for persons and their right to make autonomous decisions; (b) beneficence, or the need to maximize benefits while reducing risks; and (c) justice, or the obligation to distribute benefits and risks equally among subjects.

Research Involving Animals

Educational psychologists often study the behavior of animals as a way of understanding principles that govern human behavior. Because animals cannot consent to becoming subjects and the research may not benefit them directly, special care must be taken to ensure their best treatment. Just as in the case of protecting human subjects, Congress has developed regulations that address the responsible use of animals in research. Two specific acts, the Animal Welfare Act (1966, 1970, 1976, 1985, and 1990) and the Health Research Extension Act (1985), have provided guidance on responsible conduct. The Animal Welfare Act concerns itself with the transportation, care, and use of animals in research. The responsibility for the regulations falls upon the U.S. Department of Agriculture (USDA) under Title 9 of the Code of Federal Regulations. Guidelines were created from work in the 1950s by the Animal Care Panel, who established a professional standard for animal care in research. This work eventually led to *The Guide for Laboratory Animal Facilities and Care*, first published in 1963 and now titled *The Guide for the Care and Use of Laboratory Animals*. The Public Health Services (PHS) requires that investigators use this guidebook when establishing and conducting an institutional program of animal research. Just as human subject research requires an approval board, animal research has a similar counterpart, the Institutional Animal Care and Use Committee (IACUC). The PHS defines animals as any live, vertebrate animals used or intended for use in research, research training, experimentation, or biological testing or for related purposes. However, educational psychologists intending to use animals for research should take the broadest view possible of the definition of an animal and

should consult with their own institution's IACUC regarding their research.

Based upon the 1985 Health Research Extension Act, PHS Policy, USDA regulations, and the *Guide for the Care and Use of Laboratory Animals*, the IACUC at each institution is empowered to administer the care and use of animals in the conduct of animal research. The IACUC has responsibilities for reviewing and approving all proposals for animal research, evaluating the animal care program of the institution, inspecting the facilities, and responding to concerns about animal care. The IACUC can also exercise independent judgment to suspend research if the guidelines for care and use of animals are not followed. Federal oversight is also provided by the Office of Laboratory Animal Welfare and the USDA.

Research must follow the federal government's Principles for the Utilization and Care of Vertebrate Animals Used in Testing, Research, and Training. These include guidelines on transportation, care, and use of animals in research; designing research that is relevant to human or animal health, advancement of knowledge, or the good of society; use of appropriate species and the minimum number needed for results; avoiding or minimizing pain, discomfort, and distress of the animals; using appropriate sedation analgesia or anesthesia; painlessly euthanizing animals that are suffering; feeding, housing, and caring for animals appropriately; using only qualified and trained personnel; and consulting the IACUC about any exceptions.

In addition, a voluntary accreditation program is administered by the Association for Assessment and Accreditation of Laboratory Animal Care, a nonprofit organization that accredits programs that pass a detailed inspection. The Office of Laboratory Animal Welfare functions by relying on an assurance or signed agreement that pledges the institution to comply with all regulations and is monitored through an annual report. The USDA uses an inspection mechanism to ensure compliance.

Conflict of Interest and Commitment

Because one of the basic elements of research is that it must be free from bias, it is essential that any possible conflicts of interest or commitment be identified and rectified where possible.

Researchers with similar interests may experience conflict with one another. For example, an educational psychologist studying a new way of identifying anxiety will be interested in the research of others on this topic but may not want to share his or her results until he or she can be protected by copyright or publication. The reviewer who is conducting similar research must recognize a conflict of interest and be certain to protect the confidentiality of the material being reviewed. Conflicts are often inevitable but, in and of themselves, may not be bad. However, the researcher must be aware of them and will need to communicate their existence in order to avoid the perception that the research is unduly affected by the conflicts.

The most common conflicts include those that result in financial gain, those that affect work commitments, and those that involve intellectual and personal matters. Financial gain comes into play when a researcher unfairly delays a competitor's research so that he or she can gain an advantage. A prime example might be giving a lower score on a grant review so that one will have a better chance of getting funding for his or her similar research. At other times financial gain may not be a negative event. According to the 1980 Bayh-Dole Act, investigators and their institutions that receive federal funding are required to disclose their funding through publications, copyrights, commercialization, and other means so that knowledge can be disseminated to the public. This often will result in financial benefit in one form or another.

Conflicts of work commitment can occur when the researcher has committed to more activities than can be easily accomplished. For example, when an individual has accepted too many grants, too many peer reviews, or too many consultantships, a conflict of work commitment may exist. When the overcommitment is due to the competing needs of grants, monitoring of those activities may be accomplished through guidelines established by the Office of Management and Budget in its Circular A-21, which has regulations on how researchers can spend their time.

Conflicts of a personal and intellectual nature affect the honest and accurate conducting, reporting, and review of research. When an investigator draws conclusions that are based primarily upon personal opinion or affiliations, bias can be ascribed to the decision-making process.

Conflicts of all types need to be disclosed to avoid perception of bias in the reporting of scientific evidence. In the 1990s the federal government developed policies through the PHS and the National Science Foundation to address the handling of various conflicts. The policies require the researcher to report

significant conflicts before research is conducted, to manage conflicts of personal gain by reducing or eliminating those conflicts, and to provide information on how the conflicts have been addressed.

Research Misconduct

The volume of research conducted yearly continues to mount. With so many researchers gathering and reporting data, it is inevitable that a small number may engage in unethical practices. Even with so much research, it has been reported by Nicholas Steneck from the ORI that misconduct in research is at or below 1 case for every 10,000 researchers. He warns, however, that the number of confirmed cases may be underestimated and that underreporting is to be expected.

Research misconduct is defined by the Office of Science and Technology Policy as fabrication, falsification, or plagiarism in proposing, performing, or reviewing research or in reporting research results. This definition is similar to earlier definitions established by PHS and the National Science Foundation. Behavior is considered to meet this definition when it is a significant departure from accepted practices, has been committed intentionally or knowingly or recklessly, and has been proven by a preponderance of evidence. Although the federal definition provides a baseline for identifying abnormal practices, it does not address all of the subtle variations that constitute unethical behavior as described earlier in this entry. Therefore, many institutions add to the definition in order to encompass a broader view of research misconduct.

When an institution receives PHS funding, it must put in place a system to address accusations of research misconduct. The mechanism must include the designation of persons who are authorized to receive and investigate allegations of misconduct. It must have procedures for an initial inquiry hearing to determine if the allegations are of merit. For those situations in which the allegations seem to be well-founded, the institutions must have a process to carry out a formal investigation to discover the truth. There must be policies authorizing those responsible for the investigation to impose sanctions or vindicate the researcher and to report the results of the investigation to the ORI. There must also be regulations to protect a whistleblower who provides information in good faith.

William Wiener

See also American Educational Research Association; Institutional Review Boards; National Assessment of Educational Progress

Further Readings

Long, P. (1991). Invention, authorship, "intellectual property" and the origin of patents: Notes toward a conceptual history. *Technology and Culture*, *32*, 847–890.

Macrina, F. L. (2005). *Scientific integrity: Text and cases in responsible conduct of research* (3rd ed.). Washington, DC: ASM Press.

McSherry, C. (2001). *Who owns academic work?* Cambridge, MA: Harvard University Press.

Steneck, N. H. (2004). *Introduction to the responsible conduct of research.* Washington, DC: Department of Health and Human Services, Office of Research Integrity.

ETHNICITY AND RACE

The effects of race and ethnicity are complex and multidimensional and have institutional and social causes. This entry provides definitions for *race* and *ethnicity* in the United States as researchers, scholars, and institutional agencies currently use the terms. Although there are both biologically derived definitions and socially constructed definitions, the biologically derived definitions are being debunked, and a more complex definition, related to *racialization* (i.e., the categorization of people by race, which leads to racial discrimination) is more widely accepted. There are still many weaknesses in how researchers collect information on race and ethnicity, specifically when relying on definitions provided by the U.S. census and assuming that these definitions are consistent. However, it may be that understanding the fluidity of these definitions can allow researchers to improve data collection and data interpretation.

Race, ethnicity, and national origin have a significant impact on the life chances or life opportunities of individuals in the United States, and in the United States, education has become the means by which it is believed people gain social mobility—not solely by what they learn in the classroom but also by the social networks, prestige, and occupational credentials they gain by attending institutions of learning. Therefore, the intersections of race, ethnicity, and national origin are more profound when the disparities in educational

achievements by race, ethnicity, and national origin are considered.

The debate over the racial and ethnic disproportion in educational achievement has traditionally been explained by Linda Darling-Hammond as a process of biological fact. Due in part to the civil rights movement, a growing number of researchers of color, and a greater focus on institutional inequalities and their effects, biologically derived explanations for disparities in educational achievement have given way to cultural deficiencies explanations. Such perspectives have been articulated by Richard Herrnstein and Charles Murray and by Dinesh D'souza. The most infamous of all these works has been the discredited Herrnstein and Murray's *The Bell Curve.* Most of these works ignore the historical and systemic social and institutional barriers that are still in place today.

Defining Ethnicity, Nationalism, and Race

Within the racialization process of the United States, Blacks are at the bottom of the racial hierarchy, Whites are at the top, and other groups are in the political process of managing their positions away from Blacks and closer to Whites. Max Weber, who equated ethnic group identification to a group's belief in a shared historical origin and common ancestry, developed the historical sociological definition of *ethnicity.* In his definition, the emphasis is on belief. An ethnic relationship is believed to be equivalent to a blood relationship among the members of the ethnic group; that is, members of an ethnic group believe they share a common ancestry with members of the same ethnicity. This definition of *ethnicity* has been broadened to include the belief that a shared culture can determine an ethnic identity. Looking at ethnic identity as a process of a shared culture makes ethnic identity a process of the present-day connections between members of an ethnic group. Specifically, it is the culture that they share today that helps to determine their shared ethnic group identity. Stephen Cornell and David Hartmann have argued that this emphasis on culture has complicated ethnic identities and made them less salient. However, even though cultural interpretations of ethnicity have grown in prominence, the belief in a common ancestral original history clearly distinguishes it from other sociocultural groups found in society, such as motorcycle riders, Goths, or even social class. Ethnicity may be less salient, but the meanings of ethnic identity are still important for members of the ethnic groups.

More complex than the definition of *ethnicity* is the definition of *race. Race* has historically been defined by the perceived visibly distinct physical characteristics shared, or believed to be shared, by members of a race, whereas *ethnicity* has been defined as a process of belonging, at least contemporarily. Race is based on what makes others physically different and therefore truly not of the same biological origins. Racial definitions are usually a clearer way of creating otherness. The biological basis or what a person views as distinct traits between individuals marks a group as genetically distinct from one's own, and ethnicity does not have the same function. According to Weber, one becomes aware of one's ethnicity when one comes into contact with a person not of one's ethnic group. However, a distinct ancestry or history does not rule out the possibility of some shared bloodline, even if distant. Race, on the other hand, denies all but the minimal possibility of shared blood relationship. It is this outright biological distinctiveness that is a part of the definition of race that allows not just for the categorization of different racial groups but also for the creation of a hierarchy based on the belief that the visible racial distinctions represent only the surface of greater, more meaningful differences among races. In other words, race has historically assumed a biological distinction, and more importantly, it has inherently assumed a biological hierarchy. Ethnicity, on the other hand, marks cultural distinctiveness. There is, of course, a cultural hierarchy that is manifested by ethnic origins. For example, English and German ethnicities are usually placed higher within these categories than Italian or Greek ethnicity. However, under the proper tutelage, "bad" or "deficient" cultural values can be overcome or improved. Biological "deficiencies" cannot.

Of course, racial categorization has been a process that has continuously shifted in its defined lines. More so today, the biological association that exists with the term *race* has been further debunked by the current biological research that has shown the biological insignificance of racial categorization or the socially constructed nature of racial categories. Michael Omi and Howard Winant have noted that this has validated the claims of many social scientists and, furthermore, allowed for social scientists to show how the definition of race is clearly tied to shifts in political and societal needs and values. However, even though the

definitions of *race* and *ethnicity* are constantly redefined over time, these identities have real emotional and social implications for all the members of society. Housing, occupational, educational, and health opportunities are all affected by racial and ethnic status.

Individuals place themselves within the hierarchy of the racial and ethnic ranks, and much of their psychological identity is based on who they are and what they believe people assume their racial and ethnic origins are. Recent studies on "color-blind racism" provide some insight as to how important race and ethnic identities are, even for individuals who do not have to consider or think about their own racial categories. For example, studies on White undergraduates have clearly shown a valuing or devaluing of other individuals based on their ethnic or racial group.

Joane Nagel argues that ethnicity is a core concept, whereas national and racial categories are aspects of ethnicity. This definition may be a more appropriate definition for understanding the racialization process outside of the United States, and it can provide buffers for new immigrant groups arriving in the United States. For example, the most important factor for Africans in determining their reception in the United States was their race. However, consider the case of West Indians who are often protected from racialization in the United States. If they are assimilated into American society, then for a portion of West Indians, the effect of racialization outweighs their more favorable West Indian ethnicity or that, for some of these immigrants, their West Indian ethnicity provides some protection from a Black racialization. Ethnicity may, therefore, provide a process undergone by new immigrant groups in the United States. Specifically, ethnicity may give members of an immigrant group the ability to integrate successfully within the racialization process, by buffering some members of the immigrant group and allowing them to maintain an ethnic identity and therefore keep themselves from being considered (racially) Black. Nationalism, or the concept of a national identity, is a theoretically complicated idea, especially in the United States where even American Indian identity is proclaimed as a racial definition instead of a national definition, and although there may be political reasons for the exclusion of a national identity among American Indians, the very lack of usage of American Indian nationhood within the United States, as well as in the rest of the Americas, makes a national identity foreign or strictly understood as the identity linked to a territorial space.

However, when the concept of national identity has been used in the social sciences, it has been closely linked to the definition of ethnic identity. Nagel, for example, has defined national identity as chauvinistic, bordering, and articulating national character through history and culture for the eventual purposes of creating statehood or nationhood. What makes Nagel's view of ethnicity so important is its ability to incorporate the global racialization process that took place during the colonization and postcolonization periods. During the colonization period, the racial and ethnic hierarchical concepts from the West were imported and incorporated throughout the world. In the United States a northern European supremacy ideology arose; in Latin America an ideology of approximating or nearing Whiteness developed; and in Africa a process of mimicking Europeanness was formed. What the global colonialization period did was to set up an internal racial or ethnic hierarchy among the colonized nations and within the individuals in these societies that has continued during the postcolonial period. This internalized hierarchy reflects the earlier colonial Western racial hierarchy, although it is bound within the specific context of a postcolonial country's historical colonial experience. The Western colonizers developed a colonial administration that provided opportunities to certain groups within the colonized nation and allowed for the members of these groups to establish themselves as superior within the internally constructed racial strata.

The racialization process becomes more complex when immigrants from Latin America are considered. In Latin America, the descendants of the Spanish *conquistadores* were placed on the top of the racial hierarchy, followed by the *mestizos, mulatos, indígenas*, and *negros*. This hierarchy remained in place throughout the postcolonial period. An interesting phenomenon of the postcolonial national identity construction is how the process has developed a pro-Western or positivist ideology that made Europe the center and the creator of modern cultural, scientific, and progressive societies. The process that determines one's racial position in Latin America may be more complex than that in the United States, but it is still one that is organized around a White or European pinnacle.

These complexities of race and ethnicity are exacerbated for Latinas/os in the United States. When Latinas/os arrive in the United States, they are confronted with the fact that they often are not considered

to be White (as they may have been in their nation of origin), and many, specifically because of their lower economic status, are considered to be more like Blacks. Immigrants who come to the United States must have resources available or a support network that allows for the migration to take place. Even the very act of migrating may be indicative of a psychological orientation that allows for greater risk taking and, possibly, greater rewards.

The U.S. Census and Race and Ethnicity

The U.S. Census relies on Statistical Policy Directive No. 15, Race and Ethnic Standards for Federal Statistics and Administrative Reporting, from the Office of Management and Budget to define *race* and *ethnicity.* Whether they use census data or collect their own data, many scholars use the U.S. Census definitions provided in the decennial census. An interesting consequence of the U.S. decennial census racial and ethnic terms and definitions is the inconsistencies that can exist from decade to decade. The consequence is not that there are changes in the examples used for the racial and ethnic categories but that the very definition of a race can change to include groups previously designated as different, as with Mulattos and Blacks, or to remove groups that were included in a former census, as with Indians (Asians) and Whites. These changing definitions over U.S. history has warranted caution on the part of researchers of race and ethnicity when comparing groups across time and has led credibility to the constructiveness paradigm of race and how racialization happens in the United States.

Critical Race Theory

Critical race theory is an important construct in education and race. Critical race theory is a perspective that has been used by academics of color to help explain, identify, and critique racist elements within the structure of society. The critical race theory framework originates from critical theory, which calls on individuals to actively engage in social critique and social change. Critical race theory attempts to transform these institutional and social elements within the classroom.

Scholars in this area have identified five key elements to applying critical race theory in the classroom: (1) the centrality of race and racism and their intersectionality with other forms of subordination in education, (2) the challenge to dominant ideology around school failure, (3) a commitment to social justice in education, (4) the centrality of experiential knowledge, and (5) the transdisciplinary perspective.

Racial Beliefs or Racism

One of the consequences of having a racial hierarchy is how it becomes justified through primordial or cultural deficiencies explanations. The justification of racial inequalities through primordial explanations relieves the system of responsibility for those who benefit from the racial status quo. Thomas Jefferson found it difficult to dismiss primordial explanations for Black inferiority even when examples to the contrary were abundant. These beliefs are not merely important when attached to political leaders, but such beliefs by individuals in society can continue to manifest racial inequalities through the society. Additionally, when primordial or cultural distinctions become part of a society's discourse, all members, including those least aware, become engulfed in its ideological premises. For teachers, especially White middle-class teachers, this unconscious or conscious belief about race can be highly detrimental to a student of color.

Gabriel Aquino

See also Cultural Deficit Model; Cultural Diversity; Culture;

Further Readings

Alba, R., Logan, J. R., & Stults, B. J. (2000). The changing neighborhood contexts of the immigrant metropolis. *Social Forces*, *79*, 587–621.

Alba, R., Logan, J. R., Stults, B. J., Marzan, G., & Zhang, W. (1999). Immigrant groups in the suburbs: A reexamination of suburbanization and spatial assimilation. *American Sociological Review*, *64*, 446–460.

Darling-Hammond, L. (1998). Unequal opportunity: Race and education. *Brookings Review*, *16*, 28–32.

Gallagher, C. A. (2003). Playing the White Card: Using Ethnic Identity to Deny Contemporary Racism. In A. W. Doane & E. Bonilla-Silva (Eds.), *White out: The continuing significance of racism* (pp. 145–158). New York: Routledge.

Herrnstein, R. J., & Murray, C. (1994). *The bell curve.* New York: Free Press.

Nagel, J. (2000). Ethnicity and sexuality. *Annual Review of Sociology*, *26*, 107–133.

Portes, A., & Rumbaut, R. G. (2001). *Legacies: The story of the immigrant second generation.* Berkeley: University of California Press.

Shumway, J. M., & Hall, G. (1996). Self-selection, earnings and Chicano migration: Differences between return and onward migrants. *International Migration Review*, *30*, 979–994.

Solórzano, D., & Ornelas, A. (2002). A critical race analysis of advance placement classes: A case of educational inequality. *Journal of Latinos in Education*, *1*, 215–229.

Tate, W. (1997). Critical race theory and education: History, theory, and implications. *Review of Research in Education*, *22*, 195–247.

Waters, M. C. (1994). Ethnic and racial identities of second-generation black immigrants in New York City. *International Migration Review*, *28*, 795–820.

Webster, Y. O. (1993). *The racialization of America*. New York: Palgrave Macmillan.

ETHNOGRAPHY

Ethnography is a term that describes a broad range of research activities, including both the process of intensive field research and the textual product representing the results. Although scholars and practitioners debate the nuances of what constitutes the method of ethnography, a hallmark of ethnographic research is long-term fieldwork within a particular environment, setting, or group of people. Ethnography has its roots in cultural anthropology and sociology but is now widely used in many areas of social science, including education.

People frequently equate ethnography with qualitative research. However, not all qualitative research is considered ethnographic. Furthermore, although ethnographers rely heavily on qualitative techniques, they incorporate a wide range of information in their work, including numeric data. This entry discusses unique aspects of ethnography, as well as issues and debates on the topic, but first, a brief history of ethnography is presented.

History

Ethnography emerged in the study of exotic cultures in the late 1800s and early 1900s, because the scientific method of hypothesis testing seemed unsuitable for studying the complexity of human societies. Ethnography grew in popularity because it emphasized listening to group members in natural settings rather than manipulating subjects under contrived experimental conditions.

Anthropologist Bronislaw Malinowski was an early pioneer in ethnography, publishing a groundbreaking study of the Trobriand Islanders of New Guinea in 1922. Almost a half-century earlier, Frank Hamilton Cushing's study of the Zuni people of the American Southwest for the Smithsonian Institution was the first example of an anthropologist living with the people he studied for an extended period of time. Sociology also relies on ethnographic methods. In the 1920s and 1930s, the Chicago School of sociology studied inner city issues firsthand, such as crime and vice, homelessness, prostitutes, and gangs. Ethnography became popular in educational research in the 1970s in studies with school-age students. Schools, classrooms, playgrounds, and teacher's lounges constitute some of the many areas of ethnographic exploration in education.

Unique Characteristics of the Ethnographic Process

As mentioned at the outset, ethnography as a process relates to a family of methods involving researchers' sustained immersion in particular social or cultural contexts. Specifically, ethnography differs from other types of research in its study design, the process of data collection, and the role of the researcher.

The process of ethnography begins with a broad research question or focus, but the emphasis evolves during the course of fieldwork. The design of ethnographic research does not follow a predetermined course. Instead, meaning is constructed through an iterative process of gathering information, refining questions, and analyzing insights throughout the researcher's extended firsthand experiences among community members.

Participant observation is a fundamental method of gathering ethnographic data. It reflects the researcher's active participation in the community's daily life, simultaneously experiencing, observing, and recording social interactions among individuals in their natural environments. Although the unwritten rule in anthropology has been for fieldwork to extend across four seasons of a yearlong engagement, studies today are often compressed into a much narrower time frame. Modern ethnography, particularly outside of anthropology, also finds the researcher engaged in community activities but perhaps not living among community members. In addition to gathering information through participant observation, ethnographers also learn from interviews, key informants, and textual and visual materials.

The ethnographer's role in pursuing an insider's perspective is not that of a detached observer of events. Instead, the ethnographer is an integral part of constructing meaning through actions in the field, personal observation, and the ethnographic path followed. The ethnographer enters the community ignorant of the group's social practices and must rely on members' willingness to share their knowledge. As a result, community members are often referred to as collaborators, colleagues, or coauthors rather than participants.

Unique Characteristics of the Ethnographic Product

In addition to aspects of the ethnographic *process* just explored, the ethnographic *product* itself possesses unique characteristics. Internal validity is the primary gauge of an ethnographic text's quality. An ethnography must reveal the researcher's knowledge construction process clearly enough that readers can evaluate the text's claims. An ethnography's value is in its ability to provide a thorough, convincing, and richly detailed presentation of the insider's view of a specific situation rather than to generalize findings to other settings. Furthermore, ethnography's goal is not reliability or replicability over repeated observations. In fact, different researchers with different perspectives could generate multiple ethnographies from a given research setting. This is not to say that any version would be acceptable; rather, any one of the many possible ethnographies would require the researcher to demonstrate the ethnographic path leading to the claims made by the text.

Issues and Debates Surrounding Ethnography

Criticisms of ethnography revolve around the perceived lack of structure surrounding the research process, subjectivity involved, and the impossibility of generalizing results to other contexts. Ethnographers would argue that these limitations reflect strengths of the approach relative to hypothesis testing research. For example, lack of structure in research design allows for a full range of possible learning to emerge rather than results limited by the constraints of a specific hypothesis. Critics also claim that ethnography lacks sufficient controls for ensuring objectivity. Again, ethnographers would counter by arguing that

no research is apolitical. They would suggest that all researchers bring their own biases and perspectives to their studies, starting with the choices of research questions and study designs. The ethnographer's unique perspective is evident in the text, allowing readers to judge its influence on the work. Finally, ethnographers contend that a lack of generalizability is the price ethnography pays for rich detail and an in-depth insider's perspective.

Clearly, ethnography does not seek the same kinds of knowledge as hypothesis testing research. Ethnography has its own unique strengths and limitations. The strengths of both hypothesis testing research and ethnography are important to the pursuit of knowledge because neither one alone is appropriate for addressing the wide range of human issues to be studied.

Angela K. Murray

See also Culture; Naturalistic Observation; Qualitative Research Methods

Further Readings

Agar, M. (1986). *Speaking of ethnography.* Beverly Hills, CA: Sage.
Atkinson, P., Coffey, A., Delamont, S., Lofland, J., & Lofland, L. (Eds.). (2001). *Handbook of ethnography.* London: Sage.
O'Reilly, K. (2005). *Ethnographic methods.* New York: Routledge.
Pole, C., & Morrison, M. (2003). *Ethnography for education.* Berkshire, UK: Open University Press.
Troman, G., Jeffrey, B., & Walford, G. (Eds.). (2005). *Studies in educational ethnography: Vol. 11. Methodological issues and practices in ethnography.* Oxford, UK: Elsevier.

EVALUATION

Evaluation can be defined as disciplined inquiry to determine the worth or merit of things, where things may include programs, products, procedures, or objects. Although there is agreement about the desirability of conducting evaluations that lead to such value judgments, there are numerous related evaluation activities and models that do or do not do so. Those have been called *pseudo-evaluations* because they lack such judgments. These less-than-formal

evaluations may be a piece of a larger set of evaluative activities, or they may be self-contained studies that do not identify the value issues embedded in what is being evaluated. In addition, it is possible that they are based on points of view that eschew making judgments of worth. These differences of opinion emerge from the complexities of defining, implementing, and reporting evaluation studies as they attempt to deal with the issues discussed in this entry.

Evaluation and Research

It may be informative to distinguish activities construed as evaluation from those usually associated with research. Although there is no simple way to differentiate between the two, there are three dimensions on which they may differ. First, an evaluation need not have as its objective the generation of "new" knowledge. Evaluation is applied; research is basic. Second, evaluations produce information that is used to make decisions or form the basis of policy. Evaluations yield information that has immediate use; research may not. Third, an evaluation should be, finally, a judgment of worth. True evaluations result in value judgments; research need not and some would say should not.

Personnel Evaluations

Personnel evaluations involve making judgments about an individual's performance as well as contributions to his or her work while providing guidance to improve the individual's performance. Personnel evaluations intend to be unbiased, moral, reasonable, sufficient, and precise assessments of an individual's work. Generally, persons are evaluated over a specified time period with a particular model or approach.

There are multiple personnel evaluation approaches and systems. Most systems, however, determine individual effectiveness and contributions to their organizations as they relate to personal goals or workload agreements. Usually, the evaluations are conducted with predetermined and agreed-upon instruments, standardized categories of professional contributions and growth, and jointly acceptable goals. Scoring rubrics are applied to each evaluation category. The results of personnel evaluations are used for decisions about promotion, salary adjustments, additional training, or other personnel matters.

Program Evaluations

Program evaluations, perhaps the most popular form of evaluation, are systematic inquiries into the value, merit, and worth of educational, medical, or training programs. Proper program evaluations include not only an estimate of the effects of the program but also judgments about materials associated with the program, possible side effects of the program, and the worthiness of the goals of the program. Program evaluations can determine the quality or effectiveness of programs and provide ways to improve the program. For the most part, this encyclopedia entry describes issues and procedures related to program evaluation because of its relative popularity.

The Special Place of Goals

Evaluation studies are set in contexts where goals, standards, outcomes, or similar targets for improvement have been identified. How the goals are articulated and the importance attached to them form the basis for judging the object being evaluated. It is generally acknowledged that defining the goals specifically and rigorously is an essential first step for defining the program or project and directing the evaluation. There are those, however, who would disagree with the statement and advocate goal-free evaluation—evaluating what has been accomplished rather than what was envisaged.

Nevertheless, goals are subject to evaluation as well. A true evaluation makes a judgment of worth about the goals themselves. That judgment moves the evaluator away from a "goal attainment" model (no matter how stated or how determined, the purpose of the evaluation is to decide whether the goals have been met), to a more sophisticated evaluative stance. If, for example, the goals are trivial, it matters not whether they are met. On the other hand, a program seeking to promote high standards and attain substantively important goals can be successful even if the goals are not completely met.

There are positive reasons as well for an evaluator to make judgments about the goals: It may be the case that the goals need to be defended. For example, the goals of subject-matter curricular areas are susceptible to criticisms about the goals. "Fuzzy" mathematics has been criticized because hypothetically students do not learn to compute well. Evolution has been criticized because it is just a "theory." On the other hand,

intelligent design has been trumpeted as a competitor to evolutionary theory. In each of these cases, an evaluator and an evaluation should have a defense about the appropriateness of the goals. The needs of students and learners should dictate the appropriateness of the goals.

The level at which the goals are articulated and to whom they apply are potential pieces of an evaluation. For example, educational institutions or other organizations may have goals that are not directly related to the students they deal with. A college or university may seek, for example, to increase its retention rate or aspire to national recognition of its programs. In achieving those goals, the institution may not have affected how well a student in that institution is educated. On the other hand, if the institution had goals related to a newly defined general education program, it should have direct effects on students. It is an important part of an evaluation to determine who the audience is for both the set of activities and the outcomes of the evaluation. Those audiences are called *stakeholders* in the evaluation literature and have various degrees of affiliation with the program and potential impacts, both direct and indirect, from the program. Also, the goals may be more appropriate for one group than for another.

If an evaluation focuses on just the stated goals of a program or project, it is likely to miss important outcomes. The search for side effects and unintended consequences, both positive and negative, should be as important a focus for an evaluation as is the judgment about the goals and the attainment of the goals. For instance, an increased emphasis on drill and practice in a mathematics curriculum could have the unintended effects of increasing or decreasing students' interest in the subject or making them more or less able to solve problems. If a civics curriculum did not include a set of goals related to civic participation, it would be an inferior curriculum no matter what else it did. Also, it might have goals that preclude active civic participation, in which case it would immediately be considered undesirable. To the extent possible, a true evaluation identifies all of the important effects despite whether or not they are articulated by the goals.

Formative and Summative Evaluation

Whereas the *goals* of an evaluation are judgments of worth or value, the *roles* of evaluation may be either formative or summative. Formative and summative distinctions typically apply to the focus and timing of the evaluation. If the evaluation is focused on the early stages of a program or project, and the intent is to provide information to the developer that can be used to change that program or project, the evaluation is a formative one. If the evaluation seeks to make a final judgment about the object, the evaluation is a summative one. Just as is true for other distinctions, whether a particular evaluation effort is formative or summative can be the subject of discussion. Michael Scriven first made the distinction between formative and summative evaluations in the late 1960s.

Federal agencies, such as the National Science Foundation, often expect both internal and external evaluators of their projects. The roles of these evaluators overlap substantially with the notions of formative and summative evaluation. Although neither is necessarily relegated to one approach, the internal evaluator attends to issues most closely related to formative aspects of an evaluation; the outside evaluator tends to focus on the summative aspects and the overall effects of the program or project.

Change Over Time

The nature of many evaluations is to show change over time. Some evaluation methodologies provide baseline data and structure the evaluation to show growth from that baseline. Perhaps too infrequently, evaluations are designed to show whether there are enduring effects as well, by including a long-term follow-up study. When the intent of educational programs is to produce growth or change, a summative evaluation should determine the amount or type of change from the beginning of the program to the end. Formative evaluations provide direction for improving the program during the time frame defined by a baseline and an end point. The summative evaluation is focused on the difference between the beginning and the end.

Needs Assessments

New projects or programs should be designed and implemented to respond to a particular set of needs for a specific set of purposes. How the needs are identified and justified are questions of a needs assessment. A *needs assessment* is an evaluative process used to determine whether proposed changes are both merited and necessary. Needs assessments determine desirable improvements of a program, curriculum, or

aspects of a professional training and materials associated with a program. Needs assessments use a variety of methods to ascertain the conditions that need to be changed. Those methods include the documentation of performance, completion of specified tasks, observations, classroom data, and examination of current research evidence. Needs assessments evaluate current practices and determine needs for future trainings or programs. One set of judgments about the effectiveness of an intervention involves examining the relationships between the initial needs assessments and the end products of the program or project.

A caveat is that needs assessments may not discover needs. Needs assessments may actually uncover symptoms of needs in the forms of programmatic problems or desirable changes sometimes described as wants. Determining what are wants and what are needs is an extraordinarily important facet of a needs assessment.

Costs and Benefits

Because most evaluations are concerned with improving something and making life better for participants, it makes sense that, to some greater or lesser extent, evaluators would adopt utilitarian views of the enterprise. To the utilitarian goal of making things better for more persons, one can add that it is often desirable to be able to choose among competing alternatives and determine not only whether the outcomes are desirable but also whether one is getting the largest effect for each unit of effort. One way to address such concerns is through a cost analysis.

Cost analyses take varied forms including cost feasibility, cost-effectiveness, cost-benefit, and cost utility. To assess all of the costs, however, it is necessary to go beyond financial concerns (what does the program cost monetarily) to attempt to measure all of the costs, including such things as psychological costs, if any, to participants. In addition, economists talk about opportunity costs, those costs that are related to what other things could have been done rather than what had been done.

Cost analyses, especially cost-benefit analyses, are not easy things to do. Despite the difficulty of assessing the costs of an educational program or intervention, costs may be easier to delineate than are the benefits of that program or intervention. Enumerating benefits can be costly and time consuming. A cost-benefit analysis is often called for by those who wish

to make the judgments of worth discussed earlier. Educational programs, for instance, can provide experiences that have both short- and long-term effects. Without proper follow-up studies, therefore, the prudent evaluator may wish not to claim that a cost-benefit analysis has been conducted.

An evaluator should not eschew cost analyses simply because they are difficult. Both the effects of the intervention and the amount of time and money devoted to the evaluation should be viewed in some sensible way in terms of whether consumers are "getting their money's worth."

Professional Organizations

Professional organizations produce and orchestrate frameworks and structures of professional practice for evaluations. These organizations function as forums where professionals discuss varied perceptions and approaches to conducting evaluations and understanding evaluation processes. Professional evaluation organizations intend to improve evaluation theory and practice. At the annual conventions of the American Evaluation Association, the American Educational Research Association, and the American Psychological Association, professionals present and discuss evaluations, evaluation models, and evaluation theories with the goal of improving practice in mind.

Professional Standards

The professional evaluation standards developed by the Joint Committee on Standards for Educational Evaluation are used to guide and structure evaluations of educational and training programs, as well as other projects and materials. The standards hold evaluators accountable for making useful and accurate conclusions of merit and worth of the objects being evaluated. Evaluation standards include utility standards, feasibility standards, propriety standards, and accuracy standards. *Utility standards* require an evaluation to be useful to the various participants, audiences, and stakeholders. The utility of an evaluation depends on having met the needs of those affected (stakeholders) by the program or product being evaluated. *Feasibility standards* require evaluations to be practical, realistic, and sensible. Evaluations, therefore, need to be conducted with only the necessary levels of assets such as time, personnel, or materials. *Propriety standards* guard the rights of those participating in the

evaluation. Rules for the study of human subjects and issues related to confidentiality guide the propriety standards. *Accuracy standards* require that conclusions about merit and worth be justified and true. Conclusions about merit and worth derive from accurate and defensible evidence collected in the evaluation.

Even though there are such things as evaluation standards, evaluators and consumers of evaluations still need to make judgments about the desirability of any of the given standards.

A Model for Evaluations

There is a plethora of evaluation models, problems, and examples. They come from most of the traditional disciplines of the social sciences, humanities, and, of course, the field of education. The National Science Foundation has published a brochure, "User-Friendly Handbook for Mixed Method Evaluations," to guide evaluators who wish to approach an evaluation from a variety of perspectives. Mixed method evaluations bridge the gap between those who view evaluations in terms of social science methods and those who view them in more ethnographic terms.

Figure 1 suggests one way to define approaches to evaluation studies. An evaluator chooses one or more of the existing evaluation models, applies his or her skills to a problem, and conducts a study within a set of constraints. The three dimensions are general. The particular labels represent, more or less, the subset of skills and problems that will be addressed in an evaluation.

General Models

1. *Traditional.* An emphasis on consistency between goals, experiences, and outcomes—goal attainment models

2. *Decision oriented.* An emphasis on collecting information from a variety of sources to provide a basis for giving information to decision makers and improving the program or product

3. *Case study.* An emphasis on using a variety of ethnographic methods to understand activities and values and on being responsive to diverse audiences

4. *Integrative.* An emphasis on identifying crucial issues, side effects, and costs and benefits and making defensible value judgments
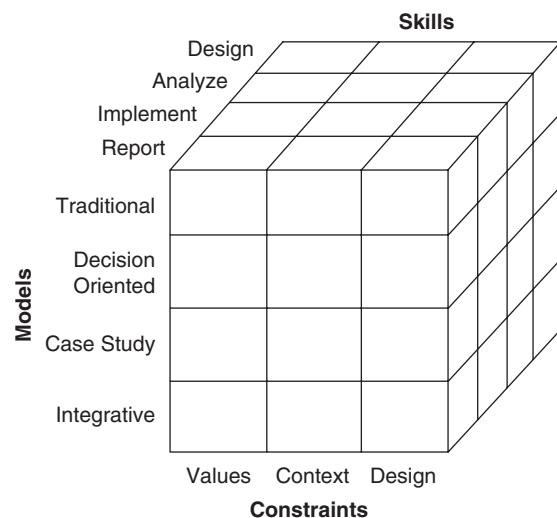
Figure 1   A heuristic for viewing evaluations

Source: Edward Kifer.

These models may incorporate different modes of inquiry. That is, a traditional evaluation focusing on goal attainment could randomly assign units to some treatment and determine the effectiveness of that treatment. A case study could include more than one setting so that comparisons could be made between the effects in one location versus those in another. The particular methods of gathering evidence—experimental, quasi-experimental, and nonexperimental designs—depend not so much on the model that is adopted as on the questions being asked and the resources available to answer them.

Presently there is a major dispute about the desirability of using methods other than randomized controlled trials or experimental field trials in either evaluation or research studies. Experimentation is called the gold standard methodology and is placed at the top of a hierarchy of possible research and evaluation methods. Experimental studies are considered the most desirable way to ascertain the impacts of programs and to determine cause and effect in such studies.

The emphasis on the superiority of experimental methods has met with considerable criticism. Responsible critics acknowledge the power of experimental methods in certain settings for specific purposes. It is generally acknowledged that a well-designed randomized controlled trial can ascertain specific effects and assign cause. However, circumstances surrounding an evaluation may make it extremely difficult to conduct

an experiment. For example, the resources available for an evaluation may not be sufficient to conduct a randomized field trial. The setting for a program may preclude the proper implementation of the program and, therefore, give erroneous results. It may not be possible to randomly assign participants to treatments, for example, to smoking and nonsmoking conditions. Critics argue that the issue is not the desirability of experimentation; rather, it is determining the extent to which it is possible, given a particular setting, to design an appropriate experiment and be able to maintain its internal validity. The fidelity of the treatment is essential.

Additionally, not all evaluations are on such a large scale that the resources are available to conduct either an experiment or a quasi-evaluation. Evaluations conducted in local settings with limited resources can still produce important and useful information. School systems and individual schools may seek information about one practice or another or about one approach versus another. The approaches and skills mentioned here may be sufficient to produce defensible answers to proposed problems. Big is not always good in evaluations.

Skills

1. *Design.* An emphasis on asking appropriate questions and gathering evidence that responds to those questions

2. *Analyze.* An emphasis on procedures and instrumentation that are congruent with the proposed evaluation design

3. *Implement.* An emphasis on data collection and data analysis; a search for consistency between this and (1) and (2)

4. *Report.* An emphasis on writing for an appropriate audience; reporting complex results in straightforward, understandable ways. Differentiating reports for varied audiences and stakeholders

There is great diversity in what an education evaluator is asked to evaluate. It can be something as short as a professional development presentation to a huge curriculum project. Much of the debate about proper methodology is set in a context of large programs with substantial evaluation budgets. However, not all (and probably a small percentage of) evaluations are of that type. Still, the evaluator has the difficult task

of making sure the evaluation contains a sound set of potentially answerable questions no matter how "small" the evaluation.

The design of the evaluation follows the questions to be asked about the object being evaluated. It is, of course, constrained by the available resources and the setting in which the evaluation occurs. Still, consistency between what is to be ascertained and how it is to be ascertained is crucial. Although there are arguments about the efficacy of ethnographic techniques in evaluation studies, few deny that how a program is implemented is a crucial factor in its success. The most obvious way to assess the "fidelity" of the treatment is through observational methods.

Technical concerns are important, too. The proper selection or construction of an evaluation tool can influence greatly the results. Sometimes there is little information to guide the evaluator. For example, although constructed response items and tasks are increasingly popular in educational assessment and evaluation, little is known about how best to design them. Technical concerns such as reliability, interrater consistency, and rubric specificity have trumped design issues. How and what is collected must be congruent with the goals of the evaluation. Such concerns suggest a substantial training need of potential evaluators— technical expertise.

A proper assessment of the evidence is the bridge between the important questions and the reporting of the results. An insightful analysis, be it a statistical or a content analysis, of the evidence provides the means to inform a variety of audiences of the purposes, goals, and effects of what is being evaluated. Multiple methods are often needed to produce the types of evidence needed for a proper analysis of the thing being evaluated.

Whereas the writer of a typical research article typically wants it published in a highly ranked, reputable journal and thus aspires to statistical sophistication demonstrating the newest techniques, the evaluator has the task of communicating important findings to a variety of audiences. Defensible evaluation publications range from a one-page glossy flyer to a well-crafted CD. In between are short reports targeting particular audiences and persons. The evaluator should choose carefully how to approach a particular audience and how sophisticated to make the presentation of the evidence. It should go without saying that the evaluator tells the same basic story to different audiences, just in simpler or more complex ways.

Constraints

1. *Values.* An emphasis on what it means to evaluate (make judgments of worth); identifying the values issues embedded in the evaluation, looking at value orientations, and discerning potential value conflicts

2. *Contexts.* An emphasis on how a setting of an evaluation impinges on what can be done effectively

3. *Design.* An emphasis on how a design of an evaluation can affect the type and strength of conclusions that might be reached

It would be interesting to know whether an evaluator has ever believed that he or she has sufficient resources to conduct the evaluation that is needed. Although resources, both financial and human, constrain what can be done by the evaluator, there are other constraints as well.

An evaluator should attempt to make judgments of worth of what is being evaluated. That means it is necessary to identify the values issues inherent in the evaluation, through the needs assessment or relevant professional literature and standards, and make sure that those issues are addressed in all aspects of the evaluation. The design must include provisions for collecting appropriate evidence so proper analyses and reporting of the values issues are a central part of the evaluation.

A good example of how contexts influence what can be done in evaluations is in the international arena and the evaluation of educational achievements. Participation by countries is often conditioned on observing traditions about what can be tested and how. On the home front, one example is what can be taught (e.g., evolution in some states). These constraints influence the very nature of comparative evaluation studies.

Finally, the evaluation question and the design determine what will be found. What goes in comes out. The evaluator must recognize that there is another day and another set of evaluations.

Corrie Orthober and Edward "Skip" Kifer

See also Assessment; Measurement

Further Readings

Levin, H. M. (1983). *Cost-effectiveness: A primer.* Newbury Park, CA: Sage.

Scriven, M. (1993). *Hard-won lessons in program evaluation.* San Francisco: Jossey-Bass.

Stufflebeam, D. L., Madaus, G. F., & Kellaghan, T. (2000). *Evaluation models: Viewpoints on educational and human services evaluation.* Norwell, MA: Kluwer Academic.

EXPERIMENTAL DESIGN

An *experimental design* is a specific strategy employed in research to answer a question of interest. In the field of educational psychology, accumulation of knowledge is based on research. For example, an educational psychologist may wish to address the question of whether or not a specific method for teaching mathematics to primary school children results in superior performance on a standardized test when contrasted with an alternative method. Another educational psychologist might conduct a study to determine whether or not noise negatively affects the reading ability of children whose school is situated in close proximity to an airport.

The three research strategies typically employed within the discipline of educational psychology are *observational research*, the *experimental method*, and *correlational research*. Whereas observational research is typically informal and subjective, the experimental method is formal and objective. Alternatively, whereas the observational method sacrifices precision for relevance, the experimental method sacrifices relevance for precision. The use of the term *precision* in defining the experimental method implies the two elements of *control* and *precise quantification*, both of which are lacking in observational research. On the other hand, the use of the term *relevance* in defining observational research reflects the fact that the latter method observes human behavior in the natural environment as opposed to studying it in under the artificial conditions associated with the laboratory experiment. A hybrid of observational and experimental research is the *field experiment*, which attempts to utilize experimental methodology to study behavior in the real world. Correlational research also attempts to provide some balance between precision and relevance, in that it can quantify the behavior of people in the real world, yet at the same time employ statistical means to impose some sort of control over the phenomenon being studied. The general subject of experimental design is most germane to research that employs the experimental and correlational methods.

Basic Definitions

The British statistician Ronald Fisher was primarily responsible for developing modern concepts of

experimental design within the framework of agricultural field experiments he conducted at the Rothamsted Experimental Station in England during the period 1919 to 1939. Among other things, Fisher introduced the concepts of randomization; blocking, which can be employed to control for extraneous variables; and factorial designs, which allow the researcher to simultaneously study the impact of multiple variables.

Regardless of the experimental design one employs, prior to conducting a study a researcher should specify a methodology that optimizes his or her ability to utilize the appropriate type of data to answer the question of interest in as efficient and precise a manner as possible. Among other things, sound experimental design involves identifying and controlling for potential sources of unwanted variability, as the latter can compromise one's ability to identify a cause-effect relationship between the variables of interest. It is important to note, however, that ethical and institutional considerations will often impose practical limitations on the type of research deemed acceptable. Consequently, the challenge to any researcher will be to design an ethically acceptable study that provides experimental control, yet at the same time has enough experiential realism such that the researcher will be able to generalize the behavior of subjects beyond the environment in which they are studied.

The most common type of hypothesis evaluated within the context of an experiment is a prediction regarding the relationship between two variables—specifically, an *independent variable* and a *dependent variable*. The independent variable is commonly referred to as the *experimental treatment*, because it represents whatever it is that distinguishes the groups employed in an experiment from one another. The number of groups employed in an experiment represents the number of *levels* of the independent variable. The measured responses of the subjects in an experiment represent the dependent variable, and if the experimenter's hypothesis is accepted, the magnitude of subjects' scores on the dependent variable should depend on the level of the independent variable to which subjects were exposed.

True Experiment Versus Natural Experiment

A distinction is often made between a *true experiment* and a *natural experiment* (which is often referred to as an *ex post facto study*). The true experiment (which

is referred to also as a *true experimental design*) is considered the gold standard of experimental design. In the simplest type of true experiment, subjects are randomly assigned to either an *experimental group* or *control group*. The experimental group only is introduced to a specific treatment manipulated by the experimenter to determine whether or not the treatment influences the behavior of subjects with respect to the response of interest, which will represent the dependent variable.

Although, because of practical or ethical considerations, it may not always be possible for a researcher to design a true experiment, the latter type of experiment optimizes one's ability to identify an existing cause-effect relationship between an independent and dependent variable, as well as to rule out one or more alternative hypotheses. The defining characteristics of a true experiment are that each of the subjects is *randomly assigned* to one of two or more groups and that the independent variable is manipulated by the experimenter. Random assignment optimizes the likelihood the groups will be equivalent to one another prior to the introduction of the experimental treatment.

At the conclusion of a true experiment, a researcher employs an inferential statistical test to determine whether or not there is a statistically significant difference between the mean scores of the two groups on the dependent variable. If a significant difference is obtained, it is likely to be due to the independent variable. It is important to note, however, that sound experimental design is a prerequisite for meaningful statistical analysis. A statistical procedure applied to a faulty experimental design will, for all practical purposes, be useless, because such a procedure is little more than an algorithm that is incapable of judging the suitability of its use.

To illustrate a true experiment, assume 100 students are randomly assigned to two different groups, each of which will be taught mathematics by one of two methods to be identified as Method A versus Method B. It will be assumed that all 100 students will be taught in the same school at approximately the same time of day (specifically, one class at 10 a.m. and the other at 11 a.m.) by the same teacher. Note that random assignment of subjects to the groups optimizes the likelihood that before being introduced to the two teaching methods (which will represent the independent variable), the two groups will be equivalent with respect to, among other things, math achievement (which will represent the dependent

variable). At the conclusion of the study, if students taught by Method A perform significantly better on a standardized math achievement test than students taught by Method B, the researcher would have a strong basis for concluding that Method A is superior to Method B for teaching mathematics.

The feature that distinguishes the true experiment from the natural experiment is that in the natural experiment, subjects cannot be randomly assigned to a group. This is because in a natural experiment, the variable that distinguishes the groups from one another is not manipulated by the experimenter but instead is a preexisting subject characteristic, such as one's gender or race. Although some researchers employ the term *independent variable* to refer to the variable that distinguishes the groups from one another in a natural experiment, others limit the use of the terms *independent* and *treatment variable* to the grouping variable employed in a true experiment. Consequently, terms such as *subject variable* or *attribute variable* may be used to designate the grouping variable employed in a natural experiment.

To illustrate a natural experiment, let us assume an educational psychologist wishes to compare the efficacy of the educational systems of two towns, Town A versus Town B, each of which comprises 1,000 students. The psychologist conducts a study in which he or she compares the scores of students in the two towns (which will represent the independent/grouping variable) on a standardized academic achievement test (which will represent the dependent variable). Results show the average score of students in Town A is significantly higher than the average score of students in Town B. Because the subjects employed in the study were not randomly assigned to the two groups (i.e., towns), the researcher will not be able to conclude that Town A has a superior educational system. Although the observed difference in academic achievement may, in fact, be due to Town A having a superior educational system, extraneous variables (such as socioeconomic status, environmental conditions, etc.) could also account for the difference. Any extraneous variable that is beyond the control of an experimenter by virtue of the fact that subjects are not randomly assigned to groups represents a potentially *confounding variable.* A confounding variable is any variable that systematically varies with the different groups. Because subjects are not randomly assigned to groups, the natural experiment is much more subject to confounding than the true experiment, and

because of the latter, if a difference between groups is obtained in a natural experiment, it does not allow a researcher to draw conclusions regarding cause and effect. For example, in the study under discussion, it is possible that the parents of students in Town A provide their children with more intellectual stimulation outside of the classroom than do the parents of students in Town B, and it is the latter variable, rather than the different educational systems, that is primarily responsible for the superior academic performance of students in Town A.

In the final analysis, the type of information one can acquire from a natural experiment is correlational in nature. Correlational information only allows a researcher to conclude a statistical association exists between the grouping variable and the dependent variable. Thus, given the design of the study that was conducted, although the researcher can conclude that higher academic achievement is associated with Town A, he or she cannot pinpoint the cause of the difference.

It should be noted the most elementary type of correlational design involves evaluating the scores of subjects on two variables to determine whether or not they are statistically associated with one another. A major goal of such research is to determine whether a subject's score on one of the variables, referred to as the *criterion variable*, can be predicted from his or her score on the other variable, referred to as the *predictor variable*. For example, a correlational design might investigate whether there is a predictive statistical relationship between the number of out-of-class activities a student participates in and a student's grade point average. More complex correlational studies can be designed that involve more than one predictor variable and one or more criterion variables.

Internal Versus External Validity

An experiment is said to have *internal validity* when its design is such that the researcher can rule out the likelihood of confounding variables. Although even with random assignment of subjects to groups there still remains the possibility of confounding, the latter is minimal, and consequently, in contrast to the natural experiment, a true experiment will typically be viewed as having internal validity. Consequently, if a significant difference on the dependent variable is observed between groups in a true experiment, a researcher will be able to argue the difference is due to the independent variable.

If the results of an experiment can be generalized, it is said to have *external validity.* More specifically, if the behavior of subjects on the dependent variable can be generalized to other persons, places, and time periods, a study will have external validity. Realistically, however, the external validity of most experiments will be limited, in that the degree to which the results of a study may be generalized will typically be limited to individuals who are comparable to the subjects employed in the study. Additionally, when a study is conducted in a laboratory or some other controlled setting, the results may not generalize to the behavior of people outside of such settings. As noted earlier, a researcher will be challenged to design a study that achieves a reasonable balance between experimental control and experiential realism. Typically, the greater the internal validity of a study, the lower the external validity, and vice versa. With respect to the latter, the true experiment is often depicted as being high in internal validity yet low in external validity, whereas observational research is depicted as being high in external validity yet low in internal validity.

Threats to Internal Validity

Donald Campbell and Julian Stanley made an important contribution to the literature on experimental design when they made a distinction between *preexperimental designs*, *quasi-experimental designs*, and *true experimental designs.* These authors noted that unlike the true experimental design, both preexperimental and quasi-experimental designs lack internal validity by virtue of the fact that subjects are not randomly assigned to experimental conditions or because a study lacks one or more control groups. A study characterized by either of the latter will not allow a researcher to effectively isolate cause and effect with respect to the relationship between the independent and dependent variables. Lack of internal validity associated with preexperimental and quasi-experimental designs can be traced primarily to the potential impact of the following potentially confounding variables: history, maturation, instrumentation, statistical regression, and mortality.

History can be defined as events other than the independent variable that occur during the period of time that elapses between a pretest and a posttest on a dependent variable. To illustrate the potential impact of history as a confounding variable, consider a hypothetical study that represents an example of a *one-group pretest-posttest design*, which is one example of a preexperimental design. Assume that 100 high school juniors are administered the SAT exam in September (which represents the pretest), after which they take a 3-month course designed to improve SAT performance. The students are then administered the SAT again in January (which represents the posttest). In the study, the SAT course can be viewed as representing the independent variable (i.e., the experimental treatment) and the difference between the pretest and posttest scores of the students as the dependent variable. If, in fact, students' January SAT scores are significantly higher than their September scores, the researcher might be tempted to attribute the increase to the SAT course. One cannot, however, rule out the possibility the increase might have been due to some other variable that was also present between September and January. For example, it is possible that during the 3 months students were enrolled in the SAT course, their classroom teachers presented lessons that were responsible for the increase in SAT performance. Consequently, without a control group composed of a comparable group of students who also took the SAT in both September and January but who did not take the SAT course, one would not be able to effectively rule out history as a potentially confounding variable.

A second threat to internal validity is *maturation*, which refers to developmental changes of a biological or psychological nature that occur within an organism as a result of the passage of time. Among other things, during the course of a study (which can be brief in duration or span a period of many years) subjects may grow stronger, become more or less agile, become fatigued, or become more or less intelligent. To illustrate the potential impact of a maturational variable on a dependent variable, assume 100 two-year-old children, who are identified, based on a pretest, as below average in visual-motor coordination, are provided with physical therapy. After 6 months of physical therapy (which represents the treatment variable), the posttest scores of the children are significantly higher than the pretest scores. Although one might surmise the physical therapy was directly responsible for the change in visual-motor coordination (which represents the dependent variable), it is entirely possible the improved performance of the children could have been due to physical maturation and that, in fact, the physical therapy had little or

nothing to do with the change in visual-motor coordination. Consequently, without a control group of children who were not provided with physical therapy, the researcher would not be able to effectively rule out maturation as a potentially confounding variable.

Instrumentation refers to inconsistencies with respect to the accuracy of the instruments employed to measure a dependent variable over a period of time. Such things as instrument malfunction or fatigue or boredom on the part of human observers charged with recording the responses of subjects are instrument-related examples that can compromise the internal validity of a study. *Statistical regression* is the phenomenon that a subject who obtains an extreme pretest score on a dependent variable will be more likely to yield a posttest score on the same variable that is closer to the mean. Consequently, in some instances, a change in a subject's posttest score can be the result of regression toward the mean rather than due to a treatment variable presented between the pretest and posttest. Another threat to internal validity is subject *mortality*. Specifically, rather than the treatment variable, differential loss of subjects in two or more groups during the course of a study may be responsible, in some instances, for a between-groups difference on a dependent variable.

Common Designs

In most cases the primary reason a design is categorized as a preexperimental design is because it lacks a control group. Yet in spite of the limitations associated with a preexperimental design, an educational researcher may occasionally employ such a design because ethical or other considerations may make it impossible or impractical to use a quasi- or true experimental design. A quasi-experimental design is more likely to be employed in educational research than is a preexperimental design, even though the internal validity of such a design is also compromised. In most instances, lack of random assignment is responsible for compromising the internal validity of a quasi-experimental design. Because many education-related issues are difficult to study through use of a true experimental design, a researcher may have no choice but to use a quasi-experimental design to investigate a hypothesis of interest. Typically, such designs will contrast the performance of two or more intact groups, such as students in different classes or towns. Use of preexisting groups in the latter type of situations does not allow a researcher to assume equivalence, as there is no reason to believe the different groups represent random samples.

A hypothetical study will be described that represents an example of a *nonequivalent control group design*, which is one type of quasi-experimental design. Assume two classes of 100 students in the same school are taught mathematics at approximately the same time of day by the same teacher. One class is taught by a method to be identified as Method A, whereas the other class is taught by a different method, to be identified as Method B. At the beginning of the school year, students in both classes are administered a standardized math achievement test that will represent a pretest measure of the dependent variable. The different teaching method each class is exposed to will represent the independent variable. At the conclusion of the school year, both classes are administered a posttest on the dependent variable. A determination with respect to whether one teaching method is superior to the other is based on the difference between the pre- and posttest scores of the two classes. At the conclusion of the study, a change score is computed for each student by computing the difference between a student's pretest and posttest scores (it will be assumed that both classes exhibit superior performance on the posttest). If the mean change score of students in the class that was taught by Method A is significantly greater than the mean change score of students who were taught by Method B, one might be tempted to conclude that Method A is superior to Method B. The latter conclusion, however, could be challenged by virtue of the fact that students were not randomly assigned to the two classes. For instance, there is the possibility that during the school year, the students in the class taught by Method A were more likely to have been exposed to conditions outside of the classroom that were conducive to increasing their skills in mathematics. If, on the other hand, students had been randomly assigned to the two classes at the beginning of the year, the likelihood of the latter would be minimal, and consequently, the design of the experiment would then conform to the requirements of a true experimental design. More specifically, use of random assignment would transform the design into a *pretest-posttest control group design*, which is one type of a true experimental design.

Another example of a quasi-experimental design is a *time-series design*, in which *multiple measures* are

obtained on a dependent variable before and after an experimental treatment. One type of time-series design is the *multiple time-series design* in which multiple measures on a dependent variable are obtained for two intact groups (i.e., groups that are formed on the basis of nonrandom assignment) before and after one of the groups is exposed to some intervention. For example, academic achievement scores might be obtained for a cohort of students for each year during a period 5 years before and 5 years after the introduction of a new curriculum in a school district. The latter scores are then compared over the same time period with the scores of what is considered to be a comparable cohort of students in another school district that did not implement the change in curriculum.

Another type of design that is often categorized as a time-series design is a *single-subject design.* This design is commonly employed with a single individual in order to demonstrate the efficacy of a behavior modification procedure on some form of maladaptive behavior. For example, in an *ABAB* design, a baseline measure of the maladaptive behavior is obtained during the initial time period, labeled A. During the second time period, labeled B, the treatment is administered and if successful, there will be a decrease in the frequency of the behavior. To confirm the treatment was responsible for the decrease in the behavior during time B, it is withdrawn during the third time period, labeled A. During the final time period, labeled B, the treatment is reintroduced in order to permanently eliminate the maladaptive behavior.

Although not categorized as time-series designs, *longitudinal* and *cross-sectional* designs are also employed in educational research to evaluate subjects' behavior over a prolonged period of time. The longitudinal design typically involves a large cohort of subjects who are repeatedly evaluated in order to determine whether or not change occurs with respect to a variable of interest with the passage of time. As an example, to determine whether or not intelligence changes during the course of one's lifetime, a cohort of individuals may be administered an intelligence test every 10 years. The same question can also be addressed through use of the cross-sectional design, which evaluates multiple cohorts comprising subjects of different age levels with respect to the variable of interest. For example, a researcher might compare the scores on an intelligence test of cohorts comprising individuals who are 10, 20, 30, 40, 50, and 60 years

of age. Because nonrandom subject mortality and historical variables can compromise the internal validity of longitudinal and cross-sectional designs, respectively, neither conforms to the requirement of a true experimental design.

A common distinction in designing an experiment is that between an *independent* and *dependent samples design.* In an independent samples design, each group comprises different subjects, whereas in a dependent samples design, the same subjects are exposed to all of the experimental treatments. A dependent samples design can also involve *matching* or *blocking* subjects. Matching subjects requires that a researcher initially identify one or more variables that are positively correlated with the dependent variable. Such a variable is referred to as a matching variable. To illustrate, consider an experiment composed of two experimental conditions in which a researcher stipulates that 10 subjects will serve in each condition. To match subjects, the researcher initially selects 10 subjects to serve in Condition 1. The researcher then identifies 10 different subjects to serve in Condition 2, with the stipulation that each subject in Condition 2 is comparable to one of the subjects in Condition 1 with respect to the matching variable (e.g., intelligence). The latter study, which will comprise 10 pairs or *blocks* of subjects, is evaluated as a dependent samples design. Although a dependent samples design is more sensitive than an independent samples design in identifying treatment differences, it is less commonly employed due to the practical problems associated with employing subjects in two or more conditions or in matching subjects.

The design of an experiment can be considerably more complex than what has been described up to this point. An example of a more complex design commonly employed in psychological research is the *factorial design*, which is able to simultaneously evaluate the impact of two or more independent variables on one or more dependent variables. A major advantage of the factorial design is it allows the researcher to identify an interaction between variables. An *interaction* is present when subjects' performance on one independent variable is not consistent across all the levels of another independent variable.

David J. Sheskin

See also Descriptive Statistics; Inferential Statistics; Scientific Method; Statistical Significance

Further Readings

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research.* Skokie, IL: Rand-McNally.

Christensen, L. B. (2000). *Experimental methodology* (8th ed.). Boston: Allyn & Bacon.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings.* Boston: Houghton Mifflin.

Fisher, R. A. (1935). *The design of experiments.* New York: Hafner.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Boston: Houghton Mifflin.

Sheskin, D. J. (2007). *Handbook of parametric and nonparametric statistical procedures* (4th ed.). Boca Raton, FL: Chapman & Hall/CRC.

EXPERT TEACHERS

A number of forces concerning educational policy have come together to raise the awareness and research concerns over the identification and preparation of *expert teachers* for America's school systems. Criteria for identifying expert teachers are being identified and studied from different perspectives, including those of teachers, principals, and students, and then examined closely as these research findings are applied to the preparation of future teachers and the induction of beginning teachers into the schools.

No Child Left Behind

There are two major components of the No Child Left Behind (NCLB) legislation that dominate current discussions related to K–12 education. The first is the closing of the achievement gap between groups of students, by focusing on standardized achievement tests of student knowledge and skills each year to provide an index of adequate yearly progress (AYP) for students in U.S. schools. The repercussions of failure to make AYP are severe and have dramatically changed curricula and teaching practices as teachers and administrators focus on preparing students to succeed in the local testing setting that varies state to state.

In 2001, President George W. Bush enacted NCLB legislation, which reauthorizes the Elementary and Secondary Education Act of 1965. These reforms were intended to increase accountability for states, school

districts, and schools; provide greater choice for parents and students; give more flexibility to states and local educational agencies in the use of federal education dollars; and put forth a greater emphasis on evidence-based teaching methods that increase student academic achievement (Executive Summary). To meet many of these intents, NCLB calls for an increase in the number of highly qualified teachers in the classroom (Section 2101). Whereas much debate in the literature exists about the definition of highly qualified teachers (also referred to as excellent teachers, effective teachers, and expert teachers), according to the NCLB legislation, a highly qualified teacher

- holds a bachelor's degree;
- is certified or licensed in his or her state;
- has demonstrated subject matter competence, as defined by the state in each of the core academic subjects he or she teaches; or
- does not have certification or licensure requirements [because they were] waived on an emergency, temporary, or provisional basis (Section 1901).

Defining Expert Teachers

Aside from the NCLB definition of expert teachers, many researchers have attempted to define the qualities of excellent teaching. In addition to determining the attributes of excellent teachers, researchers have examined the extent to which preparation helps teachers become excellent teachers as measured by student achievement. For instance, researchers have found that content preparation is positively related to student achievement in a curvilinear manner in which a threshold exists beyond which preparation is not significant in predicting student achievement. However, other research has indicated that for higher-level courses (usually in high schools), the amount of content training a teacher has had becomes very important. Although teachers often are required to be tested by their state to indicate competency, studies have reported more frequently that course-taking is a better indicator of teacher excellence than are teachers' test scores alone. This is probably because teacher assessments have not been adequately designed to accurately gather authentic measures of critical thinking, problem representation, pedagogy, and actual teaching practices.

The debate regarding the assessment of teacher effectiveness is whether coursework in education is more important than subject-matter coursework

completed in arts and sciences and related areas. Several studies have demonstrated that, indeed, education courses are more critical to teaching than are subject-matter courses. A different perspective places value on both pedagogical knowledge, as gained from education courses, and subject-matter knowledge where an interaction between the two exists. Finally, at the other end of the continuum are studies that have found that students demonstrate greater achievement when taught by teachers who are certified in their field, who have earned a master's degree, or who are enrolled in graduate courses. However, when examining this research, it is important to remember that the schools in which students are often failing are the schools in low-income neighborhoods where multiple languages are spoken and where a larger proportion of teachers are less qualified, less experienced, and more often teaching out of their certification area. As research has shown, many qualified teachers teach in schools where students are from a higher socioeconomic status. Thus, when examining teacher effectiveness, one must account for the complex relationship between teacher qualification and experience and student socioeconomic status.

The argument that effective teachers are those who are certified in their field, have a master's degree, or are enrolled in graduate studies follows the line of research about experts in general. Literature on experts indicates that experts have more domain-specific knowledge, the ability to form more accurate and better problem representations, the capacity to recognize problem-types, and the skill to be forward-working and top-down processors. Moreover, the research on expertise demonstrates that experts self-regulate and, as a result, are more flexible and can change representations faster. They also impose automaticity of lower-level cognitive processes, which frees memory capacity for more complex information. Finally, experts are experts in a specific domain; they are not experts in everything.

Researchers in the field indicate that defining an excellent teacher is difficult for a number of reasons. First, unlike fields with competitions or tournaments that identify an expert (e.g., chess and bridge), there exists no such process in education. The closest objective measurement is student performance on standardized achievement tests. Yet, this measurement of expertise is confounded by the entry characteristics of the students when the teacher receives the students assigned to him or her, not to mention the issue of

socioeconomic status and other factors confounded with test scores.

National Board-Certified Teachers

A close approximation to an objective measurement of teacher expertise is board certification, created by the National Board of Professional Teacher Standards (NBPTS). Research from this elaborate and multifaceted process found that there were 13 prototypical features of expert teachers. The features are as follows:

1. Better use of knowledge
2. Extensive pedagogical content knowledge, including deep representations of subject-matter knowledge
3. Better problem-solving strategies
4. Better adaptation and modification of goals for diverse learners and better skills for improvisation
5. Better decision making
6. More challenging objectives
7. Better classroom climate
8. Better perception of classroom events and better ability to read the cues from students
9. Greater sensitivity to context
10. Better monitoring of learning and providing feedback to students
11. More frequent testing of hypotheses
12. Greater respect for students
13. Display of more passion for teaching

These NBPTS attributes of expert teachers are consistent with the attributes of expert teachers identified by researchers in the field.

A Qualified Teacher in Every Classroom

According to NCLB regulations, the subject-matter competence for elementary teachers is defined by a state test of the basic elementary school curriculum and the state's high, objective, uniform, state standard of evaluation (HOUSSE). Secondary teachers responsible for core academic subjects, such as English, reading or language arts, mathematics, science, history, civics and government, geography, economics, the

arts, and foreign languages, must demonstrate competence in the content area(s) they are teaching. For secondary teachers, subject-matter competence must be met in each core academic subject they teach. This can be met by passing a state test in each subject or by completing an academic major, graduate degree, coursework equivalent to an undergraduate degree, or advanced certification. This competence can also be met by completing the state's HOUSSE in that subject.

The criteria for the HOUSSE (see the NCLB Toolkit)

- are set by the state for grade-appropriate academic subject-matter knowledge and teaching skills;
- are aligned with challenging state academic content standards and student achievement standards and developed in consultation with core content specialists, teachers, principals, and school administrators;
- must provide objective, coherent information about the teacher's attainment of core content knowledge in the academic subjects in which a teacher teaches;
- are applied uniformly to all teachers in the same academic subject and the same grade level throughout the state;
- take into consideration, but are not based primarily on, the time a teacher has been teaching the academic subject; and
- are made available to the public upon request.

Beyond Pedagogy

Beyond general pedagogical knowledge and pedagogical content knowledge, other characteristics have been deemed important in excellent teaching. For instance, studies of the many kinds of knowledge attributed to teachers' effectiveness have revealed the following qualities and qualifications: general academic and verbal ability, subject-matter knowledge, knowledge about teaching and learning, teaching experience, and the set of qualifications measured by teacher certification in addition to traits such as enthusiasm, perseverance, flexibility, and concern for children. Studies also show that effective primary teachers utilize more engaging activities, demonstrate greater enthusiasm about the subject matter, hold higher expectations, praise student achievements, and encourage self-regulation. Studies have also found that being child-centered, being able to motivate students, showing empathy, and demonstrating consistency could increase the success of teachers.

Quantitative Definitions to Expertise in Teaching

Two quantitative approaches have been used to identify and define expertise in teaching. *Value-added modeling* is based on a statistical procedure that was used to analyze data related to the characteristics of a group of Tennessee teachers who consistently demonstrated student learning.

Growth curve modeling employs a regression analysis to predict student learning over three points in time and then compares predicted student achievement levels with actual student achievement. Those teachers who produced higher levels of student achievement than predicted were examined for characteristics that contributed to expertise in teaching, as measured by student achievement.

Both of these approaches are valuable tools researchers can use to examine patterns of teacher expertise while also attempting to account for student characteristics.

Elementary and Secondary Expert Teachers

Researchers have also distinguished between expert teachers in the elementary versus secondary grades because of the inherent differences between the structures of many of these typical classrooms. Specifically, teachers in the elementary grades are often responsible for the total curriculum, except for special areas such as music, art, and physical education. On the other hand, a secondary teacher may be responsible only for a specific content area although there may be different courses assigned within one general area, such as mathematics, algebra, and geometry. Teachers at the primary level may be teaching a group of 25 students for the entire day, whereas secondary teachers may be teaching five different classes of 25 students a day. Differences in classroom management, discipline as well as transitioning from one topic to another, will vary with the age of the students and the content areas. Relationships between the teacher and the students will also vary due to the time together in the classroom and the nature of the relationships developed. Although there are clear differences in the issues faced by teachers of second versus tenth graders, there are also common elements of successful teaching practices shared across these two different settings.

The Teachers for a New Era Initiative

In 2001 the Carnegie Corporation, along with additional funding from the Annenberg and Ford foundations, launched the Teachers for a New Era (TNE) initiative, which focused on the preparation of excellent teachers based on three fundamental design principles: evidence-based decision making, the engagement of arts and sciences faculty in the preparation of teachers, and the preparation of teachers through clinical practice and a systematic induction process for beginning teachers in their first teaching assignments.

Eleven colleges and universities across the United States were funded to follow these three principles and provide evidence of the development of excellent teacher preparation programs to be replicated and extended. The criteria for demonstrating the preparation of expert teachers is demonstrable pupil learning, that is, the evidence that the teacher has consistently demonstrated that his or her students learn, as measured by a valid and reliable measure of academic achievement. Although this national initiative holds great promise for establishing further evidence on the preparation of expert teachers, as measured by their impact on their students' learning, it is too early to judge. Changes in curriculum in preparing teachers were initiated only several years ago and the effects of these changes will not be available until these new teachers of TNE programs have completed their training and been given ample time to refine their knowledge, skills, and dispositions as they directly influence student learning for several more years. However, this initiative holds great promise because of the consistency of the application of the design principles and the focus on demonstrable pupil learning as the criteria for determining expertise in teaching.

International Perspectives

Studies have indicated that definitions of excellent teaching have been problematic because excellent teaching is contextually and culturally defined. For instance, in the United States, participatory learning is considered a strategy of excellent learning. In China, however, where students are deferent to their teachers, participatory learning is considered to be disrespectful. Additionally, in some countries, such as Taiwan, where texts are selected and teachers are given one national curriculum, teachers' expertise in teaching and learning is very difficult to measure or evaluate. Similarly, several studies have pointed out that excellent teachers in urban schools must possess domain knowledge in urban culture, urban political economy, community, and social service support networks.

Expert Teachers at the College Level

While student performance in high-stakes testing often determines the effectiveness of teachers at the K–12 level, student ratings have a significant impact on the evaluation of teachers and their effectiveness in higher education. For numerous years, students have been asked to identify characteristics of effective teachers. Effective teachers at the college level have been studied since the 1950s, and the characteristics have differed minimally. One study found that five characteristics of effective teachers were frequently cited as follows: capture students' attention, stress important material, make good use of examples and illustrations, inspire class confidence in their knowledge of the subject, and provide clear explanations. Another study of undergraduates indicated that a thorough knowledge of content matter, as well as preparedness, organization, enthusiasm, and student-orientedness, were critical in teaching expertise.

Scott W. Brown and Paula Johnson

See also Effective Teaching, Characteristics of; Teaching Strategies

Further Readings

Berliner, D. C. (1986). In pursuit of the expert pedagogue. *Educational Researcher*, *15*(7), 5–13.

Berliner, D. C. (2001). Learning about and learning from expert teachers. *International Journal of Educational Research*, *35*, 463–482.

Borko, H., & Livingston, C. (1989). Cognition and improvisation: Differences in mathematics instruction by expert and novice teachers. *American Educational Research Journal*, *26*(4), 473–498.

Cochran-Smith, M. (2003). Teaching quality matters. *Journal of Teacher Education*, *54*(2), 95–98.

Darling-Hammond, L., & Bransford, J. (2005). *Preparing teachers for a changing world.* San Francisco: Jossey-Bass.

No Child Left Behind Act of 2001, P.L. No. 107–110, 115 Stat. 1425. (2002). Retrieved November 1, 2006, from http://www.ed.gov/policy/elsec/leg/esea02/index.html

Oakes, J., Franke, M. L., Quartz, K. H., & Rogers, J. (2002). Research for high-quality urban teaching: Defining it, developing it, assessing it. *Journal of Teacher Education*, *53*(3), 228–234.

Rice, J. K. (2003). *Teacher quality: Understanding the effectiveness of teacher attributes.* Washington, DC: Economic Policy Institute.

Sternberg, R. J., & Horvath, J. A. (1995). A prototype view of expert teaching. *Educational Researcher*, *24*(6), 9–17.

Zumwalt, K., & Craig, E. (2005). Teachers' characteristics: Research on the indicators of quality. In M. Cochran-Smith & Kenneth M. Zeichner (Eds.), *Studying teacher education: The report of the AERA Panel on Research and Teacher Education* (pp. 157–260). Mahwah, NJ: Lawrence Erlbaum.

EXPLICIT MEMORY

Explicit memory refers to the consciously mediated, deliberate attempt to retrieve a memory from the past. Explicit memory is critically important in educational psychology because the most conceptual learning by students and testing of knowledge by teachers engage explicit memory encoding and retrieval processes. In this entry, the major factors influencing explicit memory that are educationally relevant are reviewed, followed by a description of the educational implications of explicit memory processes.

Factors Affecting Explicit Memory

Explicit memory is generally measured by one of two types of memory tasks: *recognition* or *recall*. Common classroom recognition tasks are multiple-choice and true-false questions; common classroom recall tasks are fill-in-the-blank questions and essays. In general, both recognition and recall are enhanced more by encoding processes that focus on concepts and meaning than they are by encoding processes that focus on perceptual attributes of information. Furthermore, both recognition and recall are affected more by manipulations that influence the amount and type of concept-based processing people perform than by manipulations that influence the amount and type of perceptual processing people perform. Together, these two facts suggest that explicit memory is most effectively enhanced when people attend to the conceptual aspects of information.

A second consideration for understanding explicit memory in the context of education is how information is studied. Two phenomena, the *spacing effect* and the *testing effect*, are particularly important. The spacing effect refers to the finding that distributing repetitions of information over time by placing other events in between the repetitions yields considerably better memory performance than repeating the same item twice in succession. The testing effect refers to the finding that people's memory improves most when they interleave memory retrieval attempts with the study of information. Thus, alternating between studying and retrieving produces faster learning of information immediately and dramatically better retention over time compared with repeated study of information with only intermittent retrieval.

Finally, there are differences between recognition and recall retrieval processes. Recognition memory tends to be strongly influenced by the alternatives that are presented along with a correct answer. Thus, if the incorrect alternatives on a recognition memory test are very similar to the correct answer, accurate recognition will require a more precise memory than if the incorrect alternatives are dissimilar to the correct answer. Recall, on the other hand, tends to be thought of as a process whereby a person uses cues he or she is given, such as a question or concept, to search for potential answers in their memory.

One perspective that provides a useful framework for thinking about the process of recall distinguishes between similarities and differences among memories. Similarity among memory traces, such as being from the same conceptual category (e.g., fruits), is useful for searching through memory in an effort to find candidate responses. Once a candidate response is located, it is useful to have information in a memory trace that distinguishes it from other traces that are similar and therefore also possible responses. In essence, difference information helps identify a candidate memory as the exact trace a person was attempting to retrieve.

Educational Implications

The features of explicit memory reviewed in this entry have implications for instructional techniques, students' study techniques, and testing. Regarding instruction, it is generally beneficial to use classroom instruction to inspire students to process information meaningfully, such that they can access their existing knowledge base and integrate newly learned information with their existing knowledge. This will allow them to notice commonalities between new concepts and their

existing knowledge base. However, it will also be useful for classroom instruction to note features that distinguish a new concept from existing concepts. Together, these two factors import the role of similarity and difference in explicit memory retrieval into the educational context, enabling improved access to new information, as well as improved ability to determine when the new information is (and is not) appropriately applied. Furthermore, it is useful to review at periodic intervals or to recommend that students take it upon themselves to review concepts at regular intervals. This takes advantage of the spacing effect, which promotes improved long-term remembering by distributing repetitions of information over time and intervening events. Third, students should be instructed to test themselves as they study. This recommendation makes use of the testing effect, where interleaved study and retrieval have been shown to promote improved retention compared with repeated studying with little testing.

Finally, the method of testing used has implications for the kinds of knowledge students may be able to assess. The instruction and study recommendations noted in this entry are largely tailored to improving recall performance (essay, fill-in-the-blank), although they will also tend to be effective at improving recognition memory performance (multiple-choice, true-false). If the assessment used is a multiple-choice test, it is important for educators to be aware of the sensitivity of recognition memory to incorrect response alternatives. In particular, if only one response alternative on a multiple-choice question is plausible or was encountered in the classroom curriculum, students will be able to answer that question accurately based on the fact that only one possible answer contains a concept that they encountered in the classroom. On the other hand, if multiple response options are plausible and were encountered in the classroom context, students will need to possess more precise knowledge of class concepts to answer that question accurately. In essence, increasing the similarity between the incorrect test alternatives and the correct answer on a multiple-choice test should allow instructors to evaluate the specificity and precision of students' knowledge of subject matter, as opposed to simply assessing whether or not they have encountered a concept in the classroom context.

Jason Arndt

See also Episodic Memory; Long-Term Memory; Memory; Short-Term Memory

Further Readings

Bahrick, H. P., Bahrick, L. E., Bahrick, A. S., & Bahrick, P. E. (1993). Maintenance of foreign language vocabulary and the spacing effect. *Psychological Science*, *4*, 316–321.

Hunt, R. R., & McDaniel, M. M. (1993). The enigma of organization and distinctiveness. *Journal of Memory and Language*, *32*, 421–445.

Karpicke, J. D., & Roediger, H. L., III. (2007). Repeated retrieval during learning is the key to long-term retention. *Journal of Memory and Language*, *57*, 151–162.

EXPLICIT TEACHING

The term *explicit teaching* refers to a method of instruction in which the teacher, who serves as a provider of knowledge, presents skills and concepts in a clear, systematic, and direct way that promotes student mastery. *Explicit teaching* refers to the type of direct, teacher-led instructional explanation a teacher employs when describing curricular content and procedures, strategy instructions, skills and concepts, and rules for memorization through verbally detailed explanations and examples. Explanations may include the teacher modeling a skill or concept, followed by the teacher guiding and providing feedback as the student practices applying the skill or concept. Students are provided many opportunities to apply, independently, the skill or concept so as to ensure their mastery and generalization of the skill. Explicit teaching methods include the teacher reviewing previous material, describing procedures, highlighting, modeling skills using cues and prompts, questioning students directly to ensure understanding, providing feedback and progress monitoring while presenting multiple opportunities for students to practice the skill to mastery, and providing opportunities for students to apply the new skill. During the modeling process, the teacher uses "think-alouds"; that is, the teacher talks aloud about what he or she is thinking. The teacher provides specific examples of the sequence broken into small steps and the mental processes that occur as the teacher thinks aloud through the procedure, thereby modeling the incremental mental steps for the student. Think-alouds provide the student with detailed sequential step-by-step information which the student uses as an example when he or she goes through the mental process of solving the problem. Explicit teaching prevents confusion on the part of the student by always describing each step and by not assuming that a student should be able to

combine steps. Explicit explanations assist the student in knowing how and when to use the skill or concept. To retain student attention and focus, instructional procedures are presented at a brisk pace.

Explicit teaching is appropriate when the student must demonstrate a high level of mastery of a task-specific strategy, when student acquisition of skills or factual content is essential before further skills or concepts can be acquired, when the student has little or no background knowledge of concepts, or when the student experienced initial failure with content. When the complexity of the task is reduced by breaking the procedure into smaller steps, student mastery improves. With small specific steps, opportunities for student error are reduced. An explicit approach is appropriate when limited time is available, mastery is essential, or the task is complex or potentially hazardous.

Reading and Mathematics Instruction

Explicit teaching can be used for teaching reading strategies, such as mastering phonemic awareness. Later, explicit teaching ensures that the student is identifying words correctly. As the student practices, the teacher can confirm that the student is practicing the skill correctly through frequent monitoring and questioning. Daily practice of automatic decoding and reading fluency ensures a high level of student proficiency by allowing the student to shift attention from decoding to comprehension. Comprehension concepts, such as cause and effect, reality and fantasy, and literal and inferential questions, can be taught using explicit instruction.

Explicit teaching focuses on student mastery and teaching subskills sequentially, usually through small group instruction. In mathematics, explicit teaching is used to demonstrate performing steps in algorithms. The teacher describes aloud what he or she is thinking at each step. The teacher leaves a completed model with the student as a referent for cuing purposes. Imitating the teacher's process, students try solving similar problems. The teacher, monitoring the student's repeated practice, ensures a mastery level of basic mathematics facts fluency and of operational procedures by the student.

At-Risk Students

Explicit teaching is effective for helping low-performing students acquire reading, spelling, and mathematics skills. Students at risk for school failure, with learning disabilities or other special needs, or with limited experience or proficiency with lesson content are most successful when taught through explicit instruction with educators maintaining a brisk lesson pace, fostering high student accuracy, and prompting students rather than telling them correct answers.

Procedure for Explicit Teaching

Explicit teaching provides a clear statement and prevents vagueness and ambiguity, thus increasing the likelihood of understanding by the student. Using explicit teaching, the teacher

1. states the goals of the lesson;

2. reviews previous learning whose mastery is a prerequisite to learning the new skill or concept;

3. presents new material in small steps with clear, detailed, sequential step-by-step instructions and explanations;

4. models each step by physically demonstrating the directions and verbally describing the mental processes;

5. provides active student practice until each step is mastered;

6. guides, monitors, and provides systematic feedback during students' initial practice;

7. checks for student understanding through detailed questioning;

8. provides students with high levels of successful independent practice until mastery level is obtained;

9. provides students with opportunities for skill application; and

10. continues weekly and monthly reviews to ascertain students' sustainability of mastery level of skill.

The emphasis of explicit teaching is on student mastery. The pedagogical processes (i.e., modeling, highlighting, feedback, review, practice, and application) provide the teacher with many opportunities to check on skill mastery by the student. The principles of explicit teaching are inherent in direct instruction, effective teaching practices, exogenous constructivism, and behavioral theory.

Paula Hartman

See also Effective Teaching, Characteristics of; Teaching Strategies

Further Readings

Bryant, D., Hartman, P., & Kim S. A. (2003). Using explicit and strategic instruction to teach division skills to students with learning disabilities. *Exceptionality*, *11*, 151–164.

Duffy, G., Roehler, L., Meloth, M., Vavrus, L., Book, C., Putnam, J., et al. (1986). The relationship between explicit verbal explanations during reading skill instruction and student awareness and achievement: A study of reading teacher effects. *Reading Research Quarterly*, *21*, 237–250.

Mercer, C., Lane, H., Jordan, L., & Allsopp, D. (1996). Empowering teachers and students with instructional choices in inclusive settings. *Remedial and Special Education*, *17*, 226–236.

Rosenshine, B. V. (1986). Synthesis of research on explicit teaching. *Educational Leadership*, *43*, 60–69.

EXTERNAL VALIDITY

External validity refers to the ability to generalize the findings of an experiment to other settings, populations, and times. For educational psychology, this ability is a vital one, because it indicates the extent to which various programs shown to improve learning in one (perhaps artificial) setting can be confidently expected to show similar improvements in other (real) classroom settings. Without high external validity, the program may not have its intended or presumed effect in the various populations of interest. This entry will describe external validity in detail, offer suggestions for increasing it, describe threats to it, and discuss the trade-off between external validity and internal validity.

When a new educational technique or program is being evaluated in an experimental trial, of necessity the evaluation involves, for example, a specific school or district or set of classrooms. In addition, it takes place with specific children, teachers, materials, and at a particular time. However, researchers and educators clearly do not wish to restrict their conclusions to those particular schools, teachers, children, and so on; they prefer to *generalize*, that is, assume that their conclusions about the program extends to other (or even any) schools, teachers, children, and the like. The issue of external validity concerns whether such a generalization is defensible and true.

Logically, such a generalization is hardly ever completely justifiable. Even if the experiment is later replicated in many classrooms or schools, and the same result is found, it remains logically possible that the next school will find different results. Although such an outcome always remains a logical possibility, its plausibility is related to the external validity features of the experiment.

Increasing External Validity

One approach to increasing a study's external validity is to ensure that it is conducted under conditions that are both realistic and similar to those in which the program will later be used. A more demanding approach is to use several exemplars (e.g., schools, districts) that differ in numerous ways in the initial or subsequent studies and observe whether the results replicate. The most demanding approach to achieving high external validity is to randomly select a relatively large number of exemplars from among those to which the results are hoped to eventually generalize. Under such a circumstance, schools or districts might be included as a random factor in the statistical analysis, which allows a statistical basis for generalizing.

External validity information is also often available after the study has concluded and the results have been analyzed. Subgroup analyses may indicate, for instance, that only girls' learning was increased by a specific mode of teaching and that the teaching mode had no impact on boys' learning. Consequently, the effects of the teaching mode on learning cannot be generalized across gender, because the effect is only present for girls and not boys. Many experiments are specifically designed to probe external validity by examining a previously evaluated treatment program with a new population, to establish if the program is effective with that new population and thus determine the limits of generality of the treatment effect.

Threats to External Validity

Because the experiment needs to deal with specific schools, teachers, or students, any specific element of it might be shown by later research to limit or threaten external validity. However, theorists have identified three such threats that are common across a wide range of studies. One is the interaction between selection and treatment, which refers to the possibility that those who choose to participate in the experiment differ from those who choose not to participate in ways that might cause the results to differ. For example, voluntary programs that take a considerable amount of time might elicit participation only from highly motivated or needy individuals; the same results might not have been found

with less motivated or needy participants. This challenge can be avoided, in part, if participation in the experiment is made as convenient and easy as possible. Another threat to external validity is that of the setting by treatment interaction. An experimental effect observed in a classroom on a college campus may or may not generalize to a workplace setting or to a social function. Thus, the effect of the treatment may depend upon the setting in which the experiment is conducted. This concern can be alleviated by conducting the experiment in multiple settings, to determine what, if any, impact the setting has on the effectiveness of the treatment. Finally, a third threat to validity involves the interaction between history and treatment. An effect observed in an experiment conducted immediately after the September 11, 2001, terrorist attacks may not have had the same potency 5 years later. However, this threat can be diminished by conducting the experiment again, to see if the relationship still holds, or by critically examining other evidence that may support the effect.

Trade-Offs Between Internal and External Validity

In practice, external validity may frequently work in opposition to internal validity, the extent to which the conclusion that it was a specific program that altered a certain outcome, as opposed to alternative causes of that outcome, approximates truth. This is because internal validity increases as control increases, especially control over these potential alternative causes, whereas external validity increases as variability, the antithesis of control or constancy, increases. Thus, more artificiality and purity of execution tend to rule out alternative explanations, but at the same time they compromise generality and realism. Although increasing internal validity often comes at the price of limited external validity, some experimental research designs are better than others at maximizing both. For example, the pretest-posttest control group design is highly internally valid, but the necessity of assessing outcomes twice may be highly artificial, compromising external validity. Designs that are relatively high on both include the posttest only control group design; many quasi-experimental designs, such as the interrupted time series and switching replications designs; and the combined modified design, for assessing palatability of interventions simultaneously with their efficacy.

Sanford L. Braver and Melinda E. Baham

See also Evaluation; Experimental Design; Internal Validity

Further Readings

Braver, S. L., & Smith, M. C. (1996). Maximizing both external and internal validity in longitudinal true experiments with voluntary treatments: The "combined modified" design. *Evaluation and Program Planning*, *19*(4), 287–300.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research.* Chicago: Rand-McNally.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design & analysis issues for field settings.* Boston: Houghton Mifflin.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Boston: Houghton Mifflin.

EXTRACURRICULAR ACTIVITIES

Extracurricular activities encompass a broad range of organized adult-sanctioned activities that fall outside the scope of the regular curriculum and usually occur during the after-school hours. They can be sponsored by either the school or the community, though most research has focused on school-based extracurricular contexts. They include such diverse contexts as sports, drama, student government, math clubs, and service activities. Despite the differences in focus, extracurricular activities share several common features. They are generally voluntary, have a regular participation schedule, are supervised by adults, include other participants, involve some structure and rules, are organized around particular competencies, and include activities that require sustained attention, opportunities for meaningful participation, and clear feedback. This entry describes why the study of extracurricular activities is important to educational psychology, different ways to define and measure activity participation, effects of extracurricular involvement, and challenges researchers face in studying these contexts.

Importance to Educational Psychology

Extracurricular activities are an integral part of the school experience for many youth. In the 1999 National Survey of America's Families, 81% of 6- to

11-year-olds and 83% of 12- to 17-year-olds reported participating in an at least one extracurricular context over the past year, with sports being the most common activity. However, although a large percentage of youth participate in extracurricular activities, many still spend a great deal of time unsupervised during the after-school hours. In fact, 40% to 50% of young people's waking time is spent in discretionary activities. In a recent national study, on average, youth reported spending 5 hours each week participating in organized activities. At any one time, around 40% of youth were not involved in any organized activities. Most youth reported spending significantly more time in unstructured leisure activities, such as hanging out with friends or watching television, than they spent in organized activity contexts.

Another reason for the interest in extracurricular activities is the potential role of such activities in promoting achievement and preventing school disengagement. Conventional wisdom holds that by participating in extracurricular activities, individuals learn skills that will help them achieve in school, for example, a strong work ethic, respect for authority, time management, and perseverance. Participation also may have academic benefits because some types of activities are structured to reinforce classroom learning and offer students the opportunity to apply academic skills in real-world contexts. Involvement in extracurricular activities may be especially important for youth who do not excel academically, offering the individual another setting to develop skills and gain recognition. Participation in extracurricular activities also may enhance students' attachment to school and reduce the likelihood of their dropping out by contributing to their sense of being a valued member of the school community. Finally, spending time in these settings may have motivational benefits. Youth report high rates of alienation and boredom in traditional curricular contexts but positive affect and motivation in organized out-of-school activities.

Demographic changes also have fueled the interest in research on the effects of extracurricular participation. An increase in parents' work commitment has resulted in large numbers of youth who are unsupervised during the after-school hours. Research has shown that the lack of adult supervision places these children at a higher risk for negative outcomes. High rates of family mobility and the disorganization of many neighborhoods have also weakened the informal supports for many youth. Participation in organized programs, like extracurricular activities, during the after-school hours can help to compensate for the absence of family and community supports.

Effects of Participation

Researchers in such diverse areas as leisure studies, sociology, sports psychology, and adolescent development have explored the effects of extracurricular involvement. Converging evidence from these literatures generally supports the hypothesis that participation has beneficial effects. There is a more extensive literature on sports than on other activity contexts, though more recent research has begun to take a nuanced approach and examine developmental outcomes across a range of activities. This research has helped to illustrate which findings are unique to sports and which findings hold more broadly across a range of activity contexts.

Much of the research has focused on the academic effects of extracurricular involvement. Several studies have linked activity participation to grades, test scores, school engagement, attendance, and educational attainment. Other work has shown a correlation between participation and indicators of psychological adjustment. For instance, extracurricular participation is related to higher self-concept and interpersonal competence and lower rates of depression and anxiety. Involvement in high school extracurricular activities also appears to have some beneficial effects for adult development, including a higher likelihood of college attendance, more favorable mental health outcomes, and increased involvement in political and civic causes. Finally, extracurricular involvement is associated with a reduction in problem behavior. For example, participation predicts staying in school, lower rates of delinquency, and lower rates of substance use. The protective benefits seem to be largest for high-risk youth.

Although most of the effects are positive, there is evidence of some potential costs to participation in certain types of activities. For example, some studies find that sports participation is related to higher levels of alcohol use. One explanation for this finding is that drinking is a part of the culture of athletics and is supported and reinforced by other peers. Another concern is the inappropriate behavior of some parents and coaches in youth sports. The potential for negative peer group dynamics, especially in less structured extracurricular contexts, also has been highlighted. Finally, some have expressed concern over the competition and the stress associated with excessive

involvement in extracurricular activities, especially among upper-middle-class families.

Explanations for Benefits

The primary goal of most studies has been to correlate extracurricular participation with individual-level outcomes. Although these studies provide evidence of the beneficial effects of activity involvement, these predictive findings provide little information about the reasons for these associations. A few scholars have begun to explore the characteristics of activity contexts that may help to explain the effects of extracurricular participation, though this research is clearly in its infancy. One explanation is that participation in extracurricular activities links adolescents to a set of similar peers. These peer groups develop an activity-based culture with shared norms and values. Participation in extracurricular activities also affords youth the opportunity to form close bonds with caring, nonfamilial adults outside of the classroom who can provide them with support, mentorship, and advice. Another explanation for the possible benefits of participation is that in these contexts, youth can explore their identity by trying new activities and clarifying their values and interests. Participating in organized activities also may help to foster initiative, or self-directed action, because these are contexts in which youth experience high motivation, challenge, and attention. Finally, the benefits of extracurricular involvement may be explained by the fact that youth can develop social, emotional, and psychological competencies that may transfer into other domains of their lives.

Student Characteristics and Extracurricular Participation

The extent to which extracurricular participation has an impact on the outcomes of development varies as a function of individual factors, such as gender, age, and ability. With the exception of sports, girls are more likely to report participating in all other types of extracurricular activities than are boys. Another common finding is the decline in participation in many organized activities as children move through adolescence. In the later grades, there are fewer slots and increasing competition for positions in many activities. The decline also likely reflects competing demands for time from school, peers, and work. An individual's ability level also affects participation.

Many schools have policies requiring students to maintain a minimum grade point average to participate in extracurricular activities. In addition, membership in some types of activities, such as music and sports, is determined by skill level.

Contextual factors also shape extracurricular participation. Access to extracurricular activities varies depending on school and neighborhood factors. In low-income communities, there are fewer resources, like athletic fields and community centers, to support extracurricular participation. In contrast, in some upwardly mobile communities, there are concerns about the overinvolvement of youth in school and community-based extracurricular activities. Parents desire to enroll their children in extracurricular activities, and their ability to manage this involvement also varies depending on time, economic resources, and their beliefs about the value of this participation.

Measuring Extracurricular Participation

The relation between extracurricular participation and development is likely to vary as a function of the type of activity, frequency of participation, breadth, and duration of involvement. A common method for assessing involvement is to ask individuals to check off on a survey whether they had participated in different activities (yes or no) and aggregate these items into a single composite measure. One assumption underlying this measurement strategy is that all activities are similar. However, the reality is that extracurricular activities offer different affordances for developing skills and relationships with adults and peers. For example, an activity like sports takes place several times a week, involves competition, is closely supervised by a coach, and has extensive opportunities for social interaction. In contrast, an activity like a foreign language club takes place less frequently, involves fewer opportunities for social interaction, and has less contact with an adult leader. Few studies have examined qualitative differences in the nature of participation across a wide range of activities, making it difficult to identify both the commonality and diversity of developmental experiences youth encounter in extracurricular activities.

Other studies have examined the relation between the total number or total time in extracurricular activities and development. This research has been guided by a zero sum model of time involvement. Greater

participation in extracurricular activities is assumed to negatively influence achievement because time spent in these activities takes away from time that can be spent on academics. A similar concern has been raised in popular media reports about overscheduled adolescents and the negative consequences for family functioning and youth well-being. A few studies have found a curvilinear relationship between activity participation and development. This work suggests that there is a threshold in which the number of activities no longer exerts positive influence on developmental outcomes. However, in contrast to the media reports of the stressed-out and overscheduled adolescent, the majority of research has shown that more participation is beneficial. More time in organization activities is likely linked to less time in either unsupervised risky activities or unproductive activities, both of which have been associated with less positive outcomes. In fact, evidence from a recent national sample of youth, shows that only a small percentage of youth (3%–6%) spend considerable time on organized activities; most youth report spending the majority of their leisure time in unstructured activities.

Another important question concerns differential associations between the breadth of activity involvement, or combination of extracurricular activities, and outcomes. Surprisingly few studies have explored this question. Limited research shows that youth in a single extracurricular activity have more favorable outcomes than youth who are not involved in any organized activities, but youth who are involved in a variety of domains have the most beneficial outcomes. Participating in a range of extracurricular contexts may be related to the most positive outcomes because different activities provide different affordances for development. Finally, there are likely to be differences in outcomes depending on the intensity or duration of involvement. Skill building and fostering relationships with adults and peers take time. The few empirical studies that have examined the effects of the duration of involvement show a positive relationship between sustained activity participation and youth development.

Limitations of Research and Future Directions

There are several gaps in the literature on extracurricular activities. Although researchers have learned a lot about the predictive effects of extracurricular involvement, there is much less information on what is happening inside these activity contexts. In fact,

few studies outside of sport psychology have measured characteristics of the activity setting. This is a critical issue as both policymakers and educators ask what about these activities matters and why. Qualitative studies that explore the contextual features of different extracurricular contexts can advance the knowledge of activity settings in important ways. The field will also benefit from longitudinal studies that are designed to evaluate theoretically based hypotheses about possible explanations for the effects of extracurricular participation. Finally, experimental studies that actually test the hypotheses about mediating factors, such as relationships with adults or peer characteristics, can help to shed light on which features of the experience matter and why some extracurricular contexts are related to more favorable patterns of development than others.

In a recent review of organized community and after-school activities for youth, the National Research Council identified eight characteristics of high-quality programs that promote positive development:

1. Physical and psychological safety

2. Appropriate structure

3. Supportive relationships

4. Opportunities for belonging

5. Positive social norms

6. Support for efficacy and mattering

7. Opportunities for skill building

8. Integration of family, school, and community efforts

This list provides an important framework for how and why different extracurricular activities like sports, student government, and drama might promote positive development. However, more research is necessary to understand what each of these features look like in different extracurricular contexts, how they interact to effect development, and which features are most important for different activities and why.

Accounting for self-selection into and out of activities is another methodological challenge. Unlike most parts of the curriculum, extracurricular participation is voluntary. As a consequence, individuals who choose to participate are often fundamentally different from youth who are not involved. Participants tend to be of higher socioeconomic status, academic ability, and have more parental support than nonparticipants. Because these factors also predict positive developmental outcomes,

failing to adjust for self-selection into activities likely overstates the benefits of extracurricular participation. In fact, in studies that have adjusted for self-selection, the strength of the relationship between extracurricular participation and developmental outcomes is reduced and, in some cases, is no longer significant.

Although there is some research on why youth select themselves into extracurricular activities, few studies have examined the factors that explain why they maintain this involvement over time. The field will benefit from longitudinal studies that begin before children start participating in extracurricular activities and track their development through adolescence and into adulthood. Person-environment fit models represent one framework for understanding why some individuals continue to participate in extracurricular activities and others drop out, especially during adolescence. Initially, youth choose to participate in activities that match their interests and talents. They will be more likely to remain in activities if the characteristics of the setting continue to match their competencies and motivation. The decline in organized activity participation during adolescence may be explained partly by the fact that some extracurricular activities fail to change to reflect the developmental needs for autonomy, respect, and belonging during this period.

Refining the measurement of extracurricular participation is another important area of future work. *Extracurricular activity* is a blanket term used to apply to any activity that is "extra" or outside of the curriculum. There has been very little discussion in the literature of the defining characteristics that are common to all extracurricular contexts, as well as the subtle and not so subtle ways organized activities differ. There are also variations in the nature of extracurricular participation in school-sponsored and community-sponsored activities, though few studies have made this distinction. Much of the previous research has focused on describing the characteristics of single activities rather than comparing organizing activities with each other or with other aspects of young people's lives (i.e., unstructured leisure, school, family, peers). This makes it difficult to disentangle whether differences in extracurricular contexts reflect the types of individuals who choose to participate in each type of activity or the characteristics of the activities themselves.

A common approach is to group different extracurricular activities together and aggregate these activities into a single item. Failing to account for differences among activities likely masks some of the effects of

extracurricular involvement on developmental outcomes. However, considering a large number of separate activities often is not realistic because of constraints of sample size. In developing measures of extracurricular participation, researchers face an ongoing tension between conceptual clarity and practical reality. The decision to use more nuanced or more global categories of extracurricular participation depends on the sample size and goals of the study. It may be more appropriate to use nuanced measures of participation when the goal is greater theoretical understanding about the characteristics of activity settings and the unique contributions of different activities to variance in outcomes of interest. However, it may not always be feasible or even practical to include such detailed measures because of the large number of survey questions and limited number of youth who participate in some extracurricular activities.

Most of the literature has focused on the effects of participation in one activity context at one point in time. However, the reality is that youth move in and out of different extracurricular activities over the course of a year and often participate in several domains at the same time. Future research should explore patterns of participation and how these different trajectories are related to indicators of adjustment. For example, research on the consequences of very high amounts of organized activity participation over time can help to inform the debate about whether involvement in extracurricular activities has become excessive for some youth. Another interesting question is what combination of extracurricular activities is associated with the most positive outcomes. Person-centered, analytic techniques can help to illustrate different configurations of activity involvement and the relation of these profiles to a range of developmental outcomes. Moreover, for many youth, participation in unstructured activities fills a much larger portion of their leisure time than participation in extracurricular contexts. Exploring the ratio of structured to unstructured activities may be one way to assess this involvement.

Another critical issue is whether the effects of extracurricular participation vary by student characteristics. To answer who benefits most from extracurricular participation, researchers need to expand the range of populations that are being investigated. More is known about extracurricular participation and development for middle-class White youth than for individuals from diverse backgrounds. Moreover, an underlying assumption guiding much of the literature is that participation has the same effect on all youth.

The effects of extracurricular participation may differ depending on the child's gender, race, socioeconomic status, risk level, and achievement, though few studies have explored this question. There also are likely to be important differences by the level of involvement and success with an activity. For example, the star of the basketball team is likely to have a different experience than a substitute who rarely plays.

Most of the previous research has examined extracurricular contexts in isolation. These activities are nested within schools and communities and are functions of the available resources and larger culture of these contexts. Ecological models assume that development is shaped by multiple overlapping spheres of influence, such as the family, school, and peer group. Examining the experience of extracurricular participation in relation to these other contexts will provide more insight into the effects of activity participation than will exploring a single domain in isolation. Extracurricular involvement may have a different effect on development depending on whether these contexts are in synchrony or conflict. An interesting question is whether a positive experience in an extracurricular context can compensate for a negative experience in another domain. Limited research suggests that high-risk and low-income youth may benefit more from extracurricular participation than other youth. This finding may reflect the fact that these youth have access to fewer of the developmental affordances in extracurricular activities in other aspects of their lives, though more research is needed to evaluate this claim. There is a need for more sophisticated research designs to identify interrelations between different contexts and examine the role of context in shaping extracurricular participation.

In conclusion, extracurricular activities are important settings in the lives of children and adolescents. Extracurricular participation has been associated with many positive academic, psychological, and behavioral outcomes. An examination of the nature and consequences of extracurricular participation is especially timely in an era of shrinking resources for schools. Extracurricular activities are often viewed as less important than academics and are one of the first items to be cut during fiscal constraints. Future research should examine the impact of student characteristics (e.g., gender, socioeconomic status), degree of participation (e.g., type of activity, intensity), and characteristics of activity setting (e.g., peer networks, relationships with adults) on development. This information is critical for designing extracurricular activities that are appropriate and tailored to the needs of youth from diverse backgrounds.

Jennifer A. Fredricks

See also Athletics; Competition; Identity Development; Motivation; Peer Influences; Risk Factors and Development

Further Readings

Eccles, J. S., & Barber, B. L. (1999). Student council, volunteering, basketball, or marching band: What kinds of extracurricular involvement matters? *Journal of Adolescent Research*, *14*, 10–43.

Eccles J. S., & Gootman, J. A. (Eds.). (2002). *Community programs to promote youth development.* Washington, DC: National Academy Press.

Eccles, J. S., & Templeton, J. (2002). Extracurricular and other after-school activities for youth. *Review of Research in Education*, *26*, 113–180.

Feldman, A. F., & Matjasko, J. L. (2005). The role of school-based extracurricular activities in adolescent development: A comprehensive review and future directions. *Review of Educational Research*, *75*, 159–210.

Holland, A., & Andre, T. (1987). Participation in extracurricular activities in secondary school: What is known, what needs to be known? *Review of Educational Research*, *57*, 437–466.

Larson, R. W. (2000). Towards a psychology of positive youth development. *American Psychologist*, *55*, 170–183.

Larson, R. W., Hansen, D. M., & Moneta, G. (2006). Differing profiles of developmental experiences across types of organized youth activities. *Developmental Psychology*, *5*, 849–863.

Mahoney, J. L., Harris, A. L., & Eccles, J. S. (2006). Organized activity participation, positive youth development, and the over-scheduling hypothesis. *Society for Research in Child Development: Social Policy Report*, *20*, 1–30.

Mahoney, J. L., Larson, R. W., & Eccles, J. S. (Eds.). (2005). *Organized activities as contexts of development: Extracurricular activities, after-school, and community programs.* Mahwah, NJ: Lawrence Erlbaum.

Marsh, H. W., & Kleitman, S. (2002). Extracurricular school activities: The good, the bad, and the non-linear. *Harvard Educational Review*, *72*, 464–514.

F

Failure is instructive. The person who really thinks learns quite as much from his failures as from his successes.

—John Dewey

FAILURE, EFFECTS OF

All people face obstacles and failures in a variety of life tasks. In academic, athletic, and business settings, failure feedback is often clear—receiving a poor grade, losing the game, or not getting a promotion. On a personal level, failures can range from the breakup of a relationship to not accomplishing a goal. For many psychologists, the question is not if an individual will experience failure, but how will he or she respond to it. Failure can have devastating consequences for an individual because it may confirm preexisting negative beliefs or require downward revisions in self-image or goals. Repeated failures may lead to feelings of helplessness and depression. Children experiencing failure in school may experience decreases in self-efficacy and self-esteem (maladaptive reactions). However, not all people experience these events as negative. Failure may motivate some by presenting challenges or opportunities to demonstrate competence in the future. They can use this information to change behavioral strategies to obtain success or establish attainable goals (adaptive responses).

There are many reasons why individuals respond differently to failure feedback. The general model used to assess reactions to failure consists of antecedent conditions and post-feedback reactions (including affective,

cognitive, and behavioral responses). Antecedent conditions include performance expectations, personality characteristics, and the investment in the task. After receiving failure feedback, individuals will likely experience some negative affect, make attributions about why they failed, and then behave in certain ways.

It is important to note that failure feedback is a relative term: For some individuals, a grade of C is an abysmal failure, whereas for others, it represents an outstanding accomplishment. Personal expectations and reactions from others influence one's perception of failure. When failure is unexpected or when others respond negatively to the outcome (e.g., by saying "How could you have failed?" "Everyone else did very well," or "This task was so easy"), the individual's response to the feedback will likely be stronger.

Although personality dimensions also affect how individuals respond to failure (e.g., those high in neuroticism or having a high need for achievement will likely respond differently than those who are low on these scales), most educational research has addressed the individual's orientation to the task. Carol Dweck and Ellen Leggett argue that motivational orientation can predict adaptive and maladaptive responses to failure feedback. An individual's motivational orientation influences how he or she will evaluate feedback based on his or her goals. Those who have learning (or mastery) goals focus on understanding, knowledge, and improvement. These individuals are motivated to

understand, and the feedback informs them if they have mastered the material. Individuals setting performance goals seek to demonstrate ability or to gain favorable judgments from others. These individuals are motivated to defend or enhance their self-worth, and therefore, feedback is something used to make social comparisons of themselves with others. Carol Midgely and colleagues developed the Patterns of Adaptive Learning Survey to assess performance orientation. This test assumes that motivational approaches are independent, such that individuals can be high on one, both, or neither dimension.

Understanding this motivational approach can help researchers understand how individuals approach tasks and how they respond to feedback. Learning-oriented individuals tend to believe that intelligence and many other abilities develop through effort (incremental theory). They seek out tests or other performance indicators to provide them with a sense of their current ability level. When they encounter failure, they attribute causality to lack of effort. Mistakes indicate areas where more effort is needed. When faced with similar tasks in the future, they will often modify their behaviors to spend more time on mastering the task. They do not show significant decreases in self-efficacy or self-esteem. Although they experience negative emotions following failure, they show more self-compassion, offering kindness and understanding toward self, gradually decreasing the negative emotions. These individuals may reevaluate their goals, the value of the task, or the importance to mastering the material, but then they tend to avoid maladaptive responses.

On the other hand, individuals focusing on performance goals believe that abilities (including intelligence) are fixed and uncontrollable (entity theory). These individuals can adopt different goals for testing situations. Those with performance-approach goals desire feedback to gain a favorable judgment, whereas those with performance-avoidance goals avoid situations where there could be a negative judgment. In both cases, these individuals view the potential for making a mistake as threatening. These individuals tend to attribute failure to ability ("I am not smart") or external causes ("The teacher hates me"). In either case, these individuals experience decreased self-efficacy and self-esteem. They also show lower levels of self-compassion, leaning toward self-pity (exaggerating their personal suffering). Performance-avoidance goals are associated with fear of failure; these people view testing situations as threatening, and this leads

to increases in worry, fear, and anxiety. For these individuals, even success can be interpreted as "not failing." Not surprisingly, those motivated by performance goals tend to avoid situations where there will be future challenges.

Teachers, coaches, friends, and relatives often desire to reduce the negative consequences of failure for an individual and motivate him or her to succeed in the future. Although it is difficult to change an individual's motivational orientation directly, some researchers suggest that changing attributional responses can lead to more adaptive behaviors. Jeffrey Noel, Donelson Forsyth, and Karl Kelley suggest attributional retraining as a first step in helping students cope with failure. This approach provides students with models of adaptive attributional responses to failure. Students are encouraged to focus on internal controllable attributions (effort) and to avoid internal stable (ability) or external attributions. Compassionate social support is also helpful when it focuses on mastering the material rather than on evaluating the person. This strategy has yielded some success but requires more research.

Receiving failure feedback generally is a negative experience for most people, although the longer-term effects vary greatly. These individual differences can be partially explained by individual's motivational orientation. Those with a learning (mastery) orientation respond to failure with more adaptive responses than those with a performance orientation. These orientations represent complex approaches to tasks that include implicit theories of the task and the self. In helping others effectively cope with failure and persist on tasks, social support is important, specifically when it encourages the person to focus on effort and other controllable attributions.

Karl Kelley

See also Learning Strategies; Motivation; Teaching Strategies

Further Readings

Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology*, *84*, 261–271.

Ames, C., & Archer, J. (1988). Achievement goals in the classroom: Students' learning strategies and motivation processes. *Journal of Educational Psychology*, *80*, 260–267.

Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist*, *41*, 1040–1048.

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, *95*, 256–273.

Midgley, C., Kaplan, A., Middleton, M., Maehr, M. L., Urdan, T., Anderman, L. H., et al. (1998). The development and validation of scales assessing students' achievement goal orientations. *Contemporary Educational Psychology*, *23*, 113–131.

Noel, J. G., Forsyth, D. R., & Kelley, K. N. (1987). Improving the performance of failing students by overcoming their self-serving attributional biases. *Basic and Applied Social Psychology*, *8*, 151–162.

FAMILY INFLUENCES

Family influences are factors, characteristics, or behaviors exhibited by the family that can bring about changes to one or more members of that family. Though there are numerous and diverse family arrangements, research has often focused on the relationship between the parent and the child. Because parents are generally considered the primary caregivers in the family, they have a great influence on their child's biological, social, emotional, and intellectual development. Depending on the context, family influences can contribute to positive outcomes for the child as well as negative outcomes.

Genetic and Environmental Influences

One important family influence factor is the fundamental fact that parents pass their genes on to their biological children. Genes can determine the child's physical characteristics (e.g., eye color, hair color, and height), how the child's body will function, and even personality. In addition, there has been evidence that suggests that genetic factors, along with environmental influences, can contribute to the etiology of a variety of mental health and physical health disorders, such as schizophrenia, clinical depression, substance abuse, and different types of cancer. For example, a parent may pass on genes that convey risk for a specific disorder, such as schizophrenia, or a biological disposition or temperament that may allow the child to become more vulnerable only in risky social environments. If a parent or relative has a history of major depression, then the child is two to three times more likely to develop this condition than those who

do not have a parent or relative with depression. For parents with a history of alcohol or substance abuse, research evidence suggests that certain genetic influences (e.g., comorbid psychiatric disorders in parents) substantially affect the likelihood that children will develop alcohol abuse or other psychological disorders as they reach adulthood. The implication for these findings is that there may be a major genetic or hereditary influence interacting with the environment that can lead to specific developmental outcomes for the child.

However, well-functioning families can protect children who are at genetic risk and circumvent processes that usually lead to negative outcomes. For example, adopted children who are currently living in dangerous communities and whose biological parents have a history of criminality are more likely to display criminal behaviors than are adopted children living in safer neighborhoods. If these at-risk children are adopted into stable and loving families who live in safer and more supportive environments, the children's risk of developing criminal or antisocial behavior is decreased significantly. Therefore, genetic vulnerabilities that are expressed only in the presence of environmental triggers, such as parenting or environmental contexts, suggest that interpersonal and family influences can also play a significant role in the child's development.

Attachment

The behaviors of parents and caregivers influence children's behaviors and personalities as early as infancy. When emotional bonds are formed between the parents and the child, this process is known as *attachment.* There are two general types of attachment: secure and insecure. Parents, particularly mothers, who form a secure attachment with their children, are available and responsive to their child's needs and establish a sense of security within the child. Though infants initially exhibit distress when separated from their parents, the infants will reconnect warmly with the parents when they return. When frightened, securely attached children will constantly seek comfort from their parents or caregivers. These children know their parent or caregiver will provide security and reassurance, so they are comfortable seeking them out in times of need. As these children grow older, they become friendly, cooperative, and self-confident, and they adjust easily to the classroom

environment and establish amicable relationships with teachers and peers.

Children with insecure attachments to their parents, that is, children who do not have the opportunity to form such a relationship in early development, can become immature, dependent, and prone to disruptive and externalizing behavior disorders later on in life. Insecure attachment can be divided into three different groups: avoidant, anxious-ambivalent, and disorganized. Parents of children who exhibit avoidant attachment are often unavailable and unresponsive to the children's needs. As children grow older, they find it difficult to allow themselves to trust or depend on others and often prefer working alone. Children with anxious-ambivalent attachment to their parents tend to be anxious in their environment, even with their parents present. These children will often become resentful and resistant when the parent initiates attention. As children with anxious-ambivalent attachment reach adolescence, they may display social and emotional difficulties because their behavior in relationships may be perceived as needy or clingy by other people. Parents of children with disorganized attachment are seen as abusive, frightening, and severely neglectful of the child's needs. These children often show a mix of avoidant and resistant behaviors. For example, the children may exhibit dazed behavior and are often confused and apprehensive when the parent is present. Disorganized attachment appears to be a significant predictor of later childhood maladaptation, in which older children and adolescents are more vulnerable to a variety of emotional, social, and moral problems, such as aggressive and antisocial behavior.

Parenting Styles

Closely related to attachment, parenting and the parent–child relationship have been implicated repeatedly as correlates of child development both in normative situations and under adverse conditions. There are three different types of parenting styles: authoritative, authoritarian, and permissive. Parents who display an *authoritative* parenting style provide a loving and supportive home environment, hold high expectations and standards for performance, explain which behaviors are or are not acceptable, enforce household rules consistently, and include children in decision making. Though these parents are very responsive to their child's needs, they are demanding as well. Children with authoritative parents are often cooperative, energetic, self-confident, and cheerful. They often perform well academically, establish strong interpersonal relationships with peers, and show self-control and concern for the rights and needs of others.

In contrast with authoritative parents, *authoritarian* parents are highly demanding and directive and enforce strict rules without explanation or compromise. They provide well-structured environments and expect their children to obey them without question or flexibility. Children who are raised by authoritarian families tend to perform moderately well in school and be uninvolved in problem behavior, but they are often unhappy, anxious, and lacking in social skills. They also are more likely to exhibit higher levels of depression than are children from authoritative families. On the other hand, *permissive* parents are highly responsive to the children's needs but are not very demanding, as there tends to be very little structure within the home environment. Permissive parents do not require mature behavior from their children, provide no consistent set of rules, and often avoid confrontation. Children from permissive families are more likely to be involved in problem behavior and perform less well academically, but they often exhibit higher self-esteem, better social skills, and lower levels of depression than children from authoritarian or authoritative families.

Family Influences as Risk and Protective Factors

Research on risk and resilience shows that families can function as direct or indirect influences on child behavior. Family characteristics or behaviors that predict psychopathology or negative child outcomes are described as risk factors. There has been research aimed at identifying representative risk factors in development of cognitive and social-emotional competence within the family context. Some of these variables include a history of maternal mental illness, high maternal anxiety, negative or few parental interactions with the child during infancy, minimal parental education, disadvantaged minority status, stressful life events, large family size, poor family relations, and parental criminal behaviors. Many times, it is not only a single risk factor in the child, family, or environment that leads to negative outcomes but a combination of these factors that contribute to the problem in the family.

Effects of the family that demonstrate positive or desired outcomes for the child are described as protective factors for the child. Although resiliency

research has focused primarily on personal attributes, recent research also examines protective factors in a social and family context. There are currently three broad sets of variables that have been found to operate as protective factors for positive child development: (1) characteristics of the individual child (e.g., temperament, cognitive ability); (2) families that provide warmth, support, and structure; and (3) the availability of external support systems. These protective factors can help children from very high risk families and environments to avoid negative outcomes (e.g., academic failure, delinquency, depression, and substance use) and also develop successful and positive life adaptation.

Family Influences as Moderator Variables

Family influences can act as *moderator variables*, such that the influence of the family can alter the effect of another condition or factor on the child, in either a positive or negative way. In situations where there are moderating factors, there is a conditional or joint effect present, which determines the direction and/or the strength of the relationship between the predictor variable and the outcome. In one common type of interaction, positive family influence functions as a buffer to shield the child from any harmful exposure to, or impact of, adversity. For example, a protective family provides shelter and security for the child in an unsafe community so the child will not be harmed. Protective factors can often be activated by adverse events, in which parents are compelled to protect their child if they perceive that their child may be at risk or in danger. In other situations, family influences may exacerbate the effects of already-negative contexts or risk factors. For example, lack of monitoring by the parents (i.e., moderator variable) may be unsafe for children residing in bad neighborhoods (i.e., negative context) because these children do not have the security they need to live safely in these neighborhoods. Therefore, the quality of monitoring by parents in risky or dangerous environments can be described as a moderator.

Individual differences in children and parents can moderate the effects of the family as a whole, suggesting a reciprocal relationship involved. Parents can change their own perceptions and behavior in relation to their children's developmental changes and also according to how their children respond. Differences in child temperament and personality can be

moderator variables of parent behavior, and these ways of relating to people and the world often change over the course of development. In addition, the quality of parenting can be moderated by ecological systems that regulate parent behavior, such as parents' marital status, the extended family, and the social systems of work, culture, or community. For example, a supportive spouse or grandparent can be a protective factor for the children during periods of stress and hardships experienced by the family, by regulating the behavior and affect of the parent. If parents experience numerous problems at work, these problems might influence their parenting style when they come home, thereby affecting the behavior of the child.

Family Influences as Mediator Variables

Whereas a moderator variable influences the strength of the interactions between variables and the outcome, *mediator variables* serve to explain the relationship between the predictor variable and outcome. Diverse family influences can serve as mediator variables that explain the child's outcome. For example, parental involvement in a child's education is a protective factor that contributes to positive child academic performance, such as high attendance and graduation rates, improved homework completion, and decreased violence and substance abuse. Parents often intervene in their child's education if they discover that their child is failing in a subject area. As the parents work with teachers and with the child in the school, the child's grades eventually improve. Thus, parental involvement in the school acts as a mediating variable, because it helps explain how the child was able to achieve positive academic results. Parental influences can potentially mediate many risk factors, including genetic disorders, low social class, or economic difficulties, in addition to academic failure. For example, in studies that explored the genetic and environmental effects in the development of resilience among children in poverty, researchers found that maternal warmth and support mediate the effects of genetic risks to adaptation in unfavorable circumstances.

However, family influences as mediator variables can also help explain undesirable or negative child outcomes. If a child's parents are physically abusive or neglectful to the child's basic needs, this might help explain why the child is experiencing academic and social-emotional difficulties. Therefore, parenting styles

and attachments may also act as mediator variables that can affect the child's psychological well-being. Families can produce all kinds of risks, resources, and opportunities for the same child over the course of development, varying from biological and societal factors to academic and economic resources.

Family influences can also have an indirect effect on the child when the influence is mediated by intervening factors and the processes the factors represent. For example, mediator variables can be features of the child, the child's nutrition, the school, the neighborhood, the community, or any other system outside of the family that influences the child's behavior. One example of this type of mediating factor is school support, which can influence both the family and the child, especially if the family has very little education and few economic resources available to assist in their child's education. School support, therefore, acts as a positive mediating variable to improve the child's academic skills even when the family is unable to help the child succeed in school (e.g., because of stressful conditions or difficult circumstances the parents are facing).

Families Influences as Sources of Risk

Because parents are an integral part of the child's lives, they can also be a potential source of threat to their child's development and well-being. These influences can be passive in nature. Although some actions may not be necessarily deliberate, family members may harm their children's well-being through neglect or incompetence. Parents who have a physical or cognitive impairment or other mental health disorders may not be able to provide normative levels of nutrition, security, or teaching for their children, or they may not be able to notice impending danger and take protective action, which can place the child at risk for negative outcomes. Socioeconomic status and functional competence of the parents are often associated with child competence and academic achievement. For example, incompetent parents may be more likely to expose their children to danger, deviant peers, or other harmful influences in the environment due to their circumstances, choices, or behaviors. Younger children, who are still exploring their environment, are especially vulnerable to these hazards and risks if parents do not pay attention to or monitor their safety.

Family influence can also be an active source of threat. One example of this type of threat is child abuse and maltreatment, in which parents inflict physical and/or psychological harm on the child, affecting the child's welfare and safety. Studies have shown that child maltreatment perpetrated by a parent is a direct threat to children, and studies suggest that maltreatment by an attachment figure inflicts more harmful physical and/or psychological effects on the child than does maltreatment by a stranger. Children who experience maltreatment are at risk of developing health and psychological problems as adults, such as substance abuse, eating disorders, and depression. Children whose parents use harsh disciplinary methods can become defiant, explosive, and unpredictable. Even those from exceptionally abusive families tend to experience emotional difficulties and low self-esteem. In addition, because children are also social learners, children observing domestic violence or parental conflict may be disturbing for younger children, affecting their sense of security within the home environment.

There has been extensive research that explores family contextual factors for early-onset substance use (e.g., alcohol, marijuana) as children reach adolescence. Besides child abuse and maltreatment stated earlier, children whose families experience extreme poverty are considered to be at higher risk for a very early onset (before age 13) of alcohol or substance use than other children. In addition to low socioeconomic status, other early family contexts that have been shown to enhance risk for early-onset substance use include being raised in a family without a biological parent, parental stress, family medical conditions, and unemployment. However, one of the most potent risk factors is living with a parent who abuses alcohol and other substances. Research suggests that parental alcoholism and substance use significantly increase a child's likelihood that the child will experience substance use in adolescence.

Divorce, Remarriage, and the Family

Parental divorce and remarriage have an influence on the child's well-being, behavior adaptation, and academic performance. Children and adolescents whose parents are divorced or in the process of getting a divorce are at a high risk of developing externalizing behavior problems, such as antisocial behavior, aggression, and noncompliance. In addition, children and adolescents may develop internalizing behavior problems, such as depression, anxiety, low self-esteem,

and withdrawn behavior. However, parental divorce may be considered a positive solution for some families, if there is an extremely high family conflict involved. Family conflicts that involve hostility and abuse lead children to develop negative emotions that can be destructive; thus, a divorce may be an appropriate solution for the family to prevent the child from witnessing or experiencing this conflict. There may be some children who feel they are caught in the middle of this conflict or even feel helpless during this stressful period. Research suggests that parental marital hostility is associated with lack of emotional regulation in children and adolescents, which can lead to behavior problems in the future.

Marital transitions can bring positive and negative life changes to the family. Although divorce leads to an increase in stressful life events, such as poverty and psychological problems in adjustment for both child and parent, it also may allow parents the opportunity for personal growth. Remarriage provides the family an additional adult as a resource for family management and child support; however, parental monitoring tends to be lower in stepfamilies and in single-parent households than in nuclear families. In the early stages of a remarriage, stepparents are often less affectionate with their stepchildren even though they spend time with them attempting to form a relationship. In addition, biological mothers often become inattentive and less affectionate and more irritable and inconsistent in discipline. It may also be difficult for children to adjust having a new parent present especially if they still desire their biological parents back together.

Children with certain individual and family protective factors may be less likely to exhibit negative emotional and behavior outcomes after a divorce or remarriage. They may possess coping strategies, strong relationships with parents, or social support, which help children to experience less stress and emotional instability. Authoritative parenting throughout the divorce or remarriage can improve child outcomes as well. Parents who are able to provide warmth and support, along with control and monitoring of their child's behavior, can lead to diminished externalizing behavior and increase social responsibility in children and adolescents. Nevertheless, there are many challenges that arise in stepfamilies as parents try to build a marital relationship in the presence of their child and the child tries to develop meaningful and secure relationships with biological parents, stepparents, and stepsiblings as well.

Family Influences in Social and Cultural Development

Parents have numerous roles and responsibilities to teach, model, and socialize with their children in order for them to learn and adapt in their social environment. Beginning early in life, most children learn that there are some things they can or should do and other things they should not do. For instance, children learn lessons about their culture's standards and expectations, such as personal hygiene, table manners, and interpersonal skills such as knowing when to say "please" or "thank you." However, they also learn from parents not to interrupt when an adult is speaking, not to turn in homework late for school, and not to hit other children. Parents therefore attempt to instill beliefs that are acceptable in their own society. They shape or model their child's behavior through rewards and punishment, and reciprocally, the child can also influence and shape the parent's behaviors through interactions with the parents.

Families also have cultural values and beliefs systems that may influence their child to adopt those same values and beliefs or even different ones. Parents from culturally and ethnically diverse backgrounds teach or expose children to various rituals and traditions that are part of their family life. Children learn to pray or meditate and to participate in culture-specific or religious rituals. Families also have their own histories and life stories that serve many functions for children, such as imparting family values or life lessons. In various ways, parents play an important part in establishing their child's social and cultural identity through both direct education and indirect provision during their child's growth and development.

In some families, for example in Latino, Native American, and Asian communities, family bonds and relationships are especially important, and extended family members often live nearby. Children who are raised in these cultures are likely to feel responsible for their family's well-being and display loyalty and respect for their elders. It is not unusual for a child to leave school when help is needed at home. School achievement is also highly valued, and most parents encourage their child to do well academically. However, there are a few cases where academic achievement may be less valued than achievement in other areas. For example, in some Polynesian families, parents expect children to excel in creative arts and dance rather than academic subjects such as reading and writing. Therefore,

families have cultural expectations that may influence what their child is learning in school.

Many families also have different attitudes, beliefs, and perceptions about their role in their child's education, which can affect how parents communicate with teachers in their school. The customs and beliefs of culturally diverse families often conflict with those of the educational system. For some immigrant families, parents have beliefs and expectations that teachers are the experts in educating children; therefore, parents may not want to interfere in their child's schooling. For other families that are close-knit, they prefer to resolve their child's academic difficulties within the family rather than with school professionals. In addition, linguistic barriers make it difficult for parents to communicate with teachers, especially if their child is having academic or social problems in the school. Addressing these concerns may potentially help schools and parents build the partnership necessary to improve the child's learning in the school.

Family influences contribute to a child's development in positive or negative ways. They may be sources of risk and adversity or sources of support and protection. The accumulation of many risk factors leads to many problems in children; however, having the child exposed to a variety of family protective factors can contribute to positive outcomes. Parents provide resources in the family, teach children values and family traditions, and allow children opportunities to develop their skills academically and socially. Thus, in direct or indirect ways, families influence their child's development across multiple domains including the child's competence, physical and psychological well-being, adjustment, and life successes.

Andy V. Pham and John S. Carlson

See also Attachment; Child Abuse; Parenting; Parenting Styles; Risk Factors and Development

Further Readings

Clarke-Stewart, A., & Dunn, J. (Eds.). (2006). *Families count: Effects on child and adolescent development.* New York: Cambridge University Press.

Davies, P. T., & Cicchetti, D. (Eds.). (2004). Family systems and developmental psychopathology [Special issue]. *Development and Psychopathology, 16,* 477–481.

Elizalde-Utnick, G. (2002). Best practices in building partnerships with families. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology* (4th ed.,

Vol. 1, pp. 413–429). Washington, DC: National Association of School Psychologists.

Epstein, J. L. (1996). Perspectives and previews on research and policy for school, family, and community partnerships. In A. Booth & J. F. Dunn (Ed.), *Family-school links: How do they affect educational outcomes?* (pp. 209–246). Mahwah, NJ: Lawrence Erlbaum.

Potvin, P., Deslandes, R., & Leclerc, D. (1999). Family characteristics as predictors of school achievement: Parental involvement as a mediator. *McGill Journal of Education, 34*(2), 135–153.

Repetti, R. L., Taylor, S. E., & Seeman, T. E. (2002). Risky families: Family social environments and the mental and physical health of offspring. *Psychological Bulletin, 128,* 330–366.

FIELD EXPERIMENTS

Field experiments are randomized interventions that take place in naturalistic settings, as opposed to research laboratories. Education experiments may take many forms. Examples include preschool readiness programs, curriculum supplements, reductions in classroom size, and alterations in the cooperative format of the classroom, as well as larger institutional interventions, such as voucher systems that allow parents to choose among schools. In each case, field experimentation involves the random assignment of students, classrooms, or schools to treatment and control conditions.

The primary purpose of experimentation is to isolate causal relationships. Random assignment ensures that exposure to the intervention bears no systematic relationship to background factors, such as students' home environment or peer influences. Field settings enable the researcher to draw causal inferences under naturalistic conditions, which enhances the external validity of the results. Field experiments strive to address four aspects of external validity: (1) How closely does the intervention resemble what will be deployed in other settings, for instance, as the result of a new policy initiative? (2) To what extent was the experimental stimulus delivered in a context that resembles the setting within which the intervention is likely to be deployed in the future? (3) How closely do the subjects in the experiment resemble those who are likely to be presented with the intervention? (4) How closely do the outcome measures resemble

the outcomes of most interest from a policy or theoretical perspective? The ideal field experiment is one that is conducted as unobtrusively as possible, using subjects and interventions that allow for generalization and outcome measures that meaningfully gauge the short- and long-term effects of the intervention.

Field experimentation is especially useful in educational environments where the intervention and setting interact in complex ways. Here, the advantages over laboratory experimentation or observational research are clear. For example, to test the influence of class size on student performance, researchers in a laboratory study might divide subjects into different sized groups for an afternoon to see how quickly the subjects learn a new skill, such as long division. A researcher conducting an observational (nonrandomized) study of the effects of class size has the ability to observe actual classroom behavior but is unable to distinguish the apparent effects of class size from other factors, such as school funding and parental involvement; even the use of multivariate statistical methods leaves open the possibility that the treatment and control groups differ systematically in unmeasured ways. A field experiment that randomly assigns students to different sized classes during elementary school has the potential benefit of isolating this one change in the students' educational environment while maintaining the advantages of unobtrusive measurement within a naturalistic setting.

In practice, field experiments often confront practical challenges. Schools and parents may be unwilling to participate in a randomized study or to adhere to a protocol that, for example, assigns some students to classes with lower student–teacher ratios. In a class size experiment, for example, parents may withdraw their children from large classes and exit the school or pressure administrators to reassign them to small classes. This problem of noncompliance may be correctable using statistical methods such as instrumental variables regression, where the instrument is assignment to the treatment group and the independent variable is whether or not the subject received the actual treatment. Noncompliance, however, not only complicates the analysis and interpretation of the results; it also undermines the statistical power of the experiment. More difficult to correct statistically are problems of attrition, as occur when subjects assigned to one experimental group are more likely than those in the other group to exit the study. A third problem is spillover, or indirect treatment of some subjects as

a function of other subjects receiving a treatment directly. For example, if improving the educational environment among children in small classes has positive educational effects on their counterparts in larger classes, a simple comparison of treatment and control groups will underestimate the effects of the intervention. Finally, field experiments are potentially susceptible to Hawthorne effects. The mere fact that an intervention is administered under the watchful eye of the evaluator may change the quality of the intervention or the manner in which participants respond to it.

The use of randomized field experiments also raises ethical considerations. Exposing students to varying interventions may be interpreted to mean that some students will be denied, for the sake of science, the best education possible. Given this ethical concern, researchers may be unable to deny access to existing services by removing pieces of an established program; instead, they must augment an existing program with new benefits and compare the effects of these changes to the baseline outcomes of the preexisting program. Sometimes resource limitations create opportunities for acceptable randomized interventions. For example, if school vouchers are in limited supply, and receiving a school voucher is determined by a lottery, researchers may discern the influence of the voucher program by comparing the educational outcomes of those who were randomly selected to receive a voucher to those who applied but were not selected.

When field experiments are impractical for logistical or ethical reasons, researchers may examine natural experiments, that is, situations in which events or institutional practices sort subjects into treatment and control groups in ways that mimic randomized assignment. For example, when a benefactor offers college scholarships to all third graders in a given school, one may examine the effects of this incentive program on graduation rates by comparing to these third graders the corresponding outcomes of second and fourth graders in the same school. Similarly, it is possible to examine the influence of kindergarten enrollment on a child's intellectual development by comparing children whose birthdays fall immediately before and after a specific cutoff date. Natural experiments permit researchers to examine the consequences of "as if" random variation. The drawback of natural experiments is that "as if" randomization is not necessarily the same as actual randomization. The researcher may be mistaken in asserting that the treatments were applied

in a random fashion, in which case, causal inferences may be biased.

Donald P. Green and Rachel Milstein Sondheimer

See also Descriptive Statistics; Experimental Design; External Validity; Internal Validity

Further Readings

Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, *91*(June), 444–455.

Cook, T. D. (2003). Why have educational evaluators chosen not to do randomized experiments? *Annals of the American Academy of Political and Social Science*, *589*, 114–149.

Cook, T. D. (2005). Emergent principles for the design, implementation, and analysis of cluster-based experiments in social science. *Annals of the American Academy of Political and Social Science*, *599*, 176–198.

Green, D. P., & Gerber, A. S. (2002). Reclaiming the experimental tradition in political science. In I. Katznelson & H. V. Milner (Eds.), *Political science: The state of the discipline* (3rd ed., pp. 805–832). New York: W. W. Norton.

Heckman, J. J., & Smith, J. A. (1995, Spring). Assessing the case for social experiments. *Journal of Economic Perspectives*, *9*, 85–110.

Mosteller, F., & Boruch, R. (Eds.). (2002). *Evidence matters: Randomized trials in education research.* Washington, DC: Brookings Institution Press.

Rosenzweig, M. R., & Wolpin, K. I. (2000, December). Natural "natural experiments" in economics. *Journal of Economic Literature*, *38*, 827–874.

FIELD INDEPENDENCE– FIELD DEPENDENCE

Field independence–field dependence (FI/D) is a dimension of cognitive style that affects the degree to which individuals rely on their own internal frames of reference in perception and performance of cognitive and social tasks. The degree of reliance on internal or external referents, in turn, affects other cognitive and social domains of behavior. The result is the portrait of a field independent person who is perceptually and cognitively analytic—able to see the trees in the forest. The same person is socially introverted, thinking his or her own thoughts rather than engaging with others. The field dependent individual perceives and thinks more holistically and can even be overwhelmed by extensive data in the field—unable to engage selective attention without help. The field dependent person relies on others for direction, help, and confirmation and thus exhibits a socially engaged interpersonal style. People develop from a relatively field dependent style as children to more field independent with age, but this general developmental trend does not have a uniform outcome, and therefore, FI/D remains a variable of individual difference within the adult population.

The descriptions of individuals falling at the extreme ends of the FI/D continuum are immediately recognizable by any teacher. Some students seem to take off on their own, cognitively speaking. These field independent learners like to work independently and even get impatient when they are placed in a group for discussion and problem solving. Their field dependent counterparts are happy when it is time for group work. They are lost on their own and, therefore, subscribe to the notion that two heads are better than one. The fact that these types of individuals are so recognizable to teachers may be one reason that the construct of FI/D remains important in educational psychology.

A second reason is that all educationalists are attracted to the idea that individual differences in students can be characterized in terms of *style* rather than solely in terms of *ability*. A style is expressed on a horizontal continuum with neither end being inherently superior to the other. Ability is expressed using a vertical metaphor, with some students on top and others at the bottom. The top is good; the bottom is not. Bottom students need remediation; top students do not. Differences in the style continuum suggest that all learners can do well if they are situated in the appropriate learning environment. In this sense, cognitive styles such as FI/D are the predecessors to the idea of multiple intelligences, which also has appealed to teachers. The utility of these conceptions of individual differences is in the promise they offer for learners and the responsibility they place on educators for structuring instruction in a way that helps learners to succeed.

Despite the conceptual attraction that the FI/D construct holds, the empirical research on this construct has been plagued by measurement problems. The original measure of FI/D in the mid-1900s was administered

individually and required the examinee to orient perceptually and physically to a field of objects as they were rotated. This cumbersome procedure was replaced by the Group Embedded Figures Test, which consists of a booklet containing line drawings in which examinees are to find hidden shapes. The rationale for the task design is that children with the field independent style will succeed in finding the hidden shapes without becoming distracted by the overall figures. The task resembles sections of intelligence tests that measure fluid ability, which is one aspect of some definitions of intelligence. Indeed, the Group Embedded Figures Test has been found to load on fluid ability when factor analyzed. Moreover, the scoring of this measure results in an overall score indicating number of correct responses; a high score means high field independence. In other words, this measure transforms the attractive, value-neutral style construct into a value-laden ability. Some connection undoubtedly exists between the field independent end of the style continuum and the field independent ability (or fluid intelligence) measured by the Group Embedded Figures Test, but research attempting to understand the role of the FI/D style in learning requires better measurement procedures.

Because of the measurement problem, some researchers have become discouraged about the utility of FI/D as a construct that can help educators understand learning. However, others see this construct as extremely salient for success in the variety of contexts that learners face today and, therefore, these researchers attempt to improve its measurement. Moreover, research results typically reveal a weak to moderate correlation between the Group Embedded Figures Test and measures of learning success under particular conditions, such as classroom second language learning. In other words, empirical results, although not dramatic, are sufficient to keep some researchers interested. In particular, researchers studying learning through technology see FI/D as an important means for understanding how individual differences intersect with technology use.

Can instruction be designed in a way that fits individual styles? Can online instruction be modified on the basis of dynamic assessment of style? Can teachers use technology to provide individual and social alternative learning environments for learners in the same class? These are some of the questions that researchers hope to address, but in order to obtain results about the success of such learning conditions,

a valid measure for assessing the FI/D style is needed. Whereas findings about individual differences may have been difficult to act on in classrooms, the range of learning options availed through technology prompts researchers to seek solutions to measurements issues.

Carol A. Chapelle

See also Cognitive and Cultural Styles; Fluid Intelligence; Individual Differences; Intelligence Tests; Multiple Intelligences

Further Readings

Abraham, R. G. (1985). Field independence–dependence and the teaching of grammar. *TESOL Quarterly*, *19*(4), 689–702.

Chapelle, C., & Green, P. (1992). Field independence/ dependence in second language acquisition research. *Language Learning*, *42*(1), 47–83.

Jonassen, D. H., & Grabowski, B. L. (1993). *Handbook of individual differences, learning, and instruction.* Hillsdale, NJ: Lawrence Erlbaum.

Witkin, H., & Goodenough, D. (1981). *Cognitive styles, essence and origins: Field dependence and field independence.* New York: International Universities Press.

FLASHBULB MEMORIES, THE NATURE OF

Psychologists interested in people's memories of life events have suggested that memories of some events may be different from others in that people are able to recall not only the events but also the *circumstances* surrounding the events and, further, that these memories may be more detailed and subject to less forgetting than others. These memories have been termed *flashbulb memories.* Two fundamental questions psychologists have posed with regard to flashbulb memories are whether they are somehow special with regard to the nature of the information stored about the event and whether they respond differently to the ravages of time than other memories. Another important question is what factor or factors contribute to the formation of a flashbulb memory. In other words, what, exactly, is necessary for individuals to form a flashbulb memory?

Flashbulb Memories and the Study of Memory in General

Psychologists interested in how individuals acquire and use knowledge have devoted much attention to the study of memory. Indeed, there have been more studies of memory than of any of the other cognitive processes that psychologists have investigated. From both a theoretical and practical standpoint, the study of memory is important and provides valuable information to psychologists and educators about the way to best present material to enhance learning, the most effective strategies to promote remembering, and an understanding of the variety of factors that often serve to limit how much individuals can remember. Much attention has been paid to how individuals of all ages encode, store, and retrieve information they have experienced. The study of memory has had a long history in psychology and remains a topic of great study and fascination today.

Psychologists have found it helpful to categorize memories into short-term and long-term memories. Short-term memories of newly presented information, such as someone's phone number, may last for several seconds but are often displaced when individuals are distracted in some way or kept from rehearsing the information. Thus, a new phone number that one has looked up and intends to dial may be lost as one answers a question posed by a child entering the room. Chances are, if the information is not familiar, for example, if this is the first time that one has attempted to think about or remember this particular phone number, the phone number that is displaced is unlikely to be remembered at a later time. It has, for all practical purposes, vanished. Long-term memories are generally thought to be more permanent than short-term memories (although all long-term memories are not necessarily permanent) and can last for several minutes up to several years. An adult's recall of the first day of kindergarten, assuming that such a memory exists, would be a good example of a long-term memory. People's ability to recall information can be very impressive, as in the case of many long-term memories, or not, as in the case of many short-term memories.

Long-term memories are generally classified into one of several types of memory. Within the broader category of long-term memory are *episodic* memories, or memories for events; *semantic* memories, or memories for general knowledge (e.g., that a kangaroo is a type of mammal); and *procedural* memories, or memories for procedures involved in various skills (e.g., in typing or playing the piano). A flashbulb memory is a type of episodic memory.

With regard to episodic memories, psychologists have been particularly interested in the numerous factors that produce either heightened or hindered recall of an event. As mentioned earlier, individuals' long-term memories can be quite impressive. What are the characteristics of an event that may make one's memory for the event particularly salient and long-lasting? Are some memories so "special" that they may be even more salient and long-lasting than others? These are questions that have intrigued not only psychologists but also individuals in all walks of life.

Studies of Flashbulb Memories

An early article by R. Brown and J. Kulik is generally thought to be the first article detailing an examination of flashbulb memories. The article is quite important, because it motivated others to begin to examine flashbulb memories in a systematic manner and because the authors presented somewhat inconsistent conceptualizations of flashbulb memories. Brown and Kulik suggested that when individuals experience an event in their life that is highly unexpected as well as consequential, they are likely to remember not only the event but also the contextual circumstances surrounding the event. So, for example, an individual might still remember the circumstances surrounding being told that her father died (such as the time of day, the activity being engaged in, where she was when given the news, what happened immediately afterward), even though several decades have passed, or upon hearing the news of the *Challenger* explosion. It should be noted that psychologists have most often studied individuals' memory of public rather than private events. Because of their nature, highly unexpected and consequential public events are likely to produce memories in a large number of people, making public memories more open to psychologists' investigations than are more personal and private flashbulb memories. Thus, psychologists know more about public than private flashbulb memories.

In an initial study conducted by Brown and Kulik, individuals were asked, in part, whether they recalled the personal circumstances of hearing the news that President John F. Kennedy had been assassinated (since Brown and Kulik's study, psychologists have investigated several other highly surprising and

consequential events, such as other assassinations, the *Challenger* explosion, the Loma Prieta earthquake, and the terrorist attacks in the United States on September 11, 2001). To have what Brown and Kulik considered a flashbulb memory, participants had to indicate that they did, indeed, remember the circumstances surrounding being made aware that Kennedy had been killed and also had to be able to provide information about at least one of six "canonical" categories of contextual circumstances (place, ongoing event, informant, affect in others, own affect, and aftermath). What is interesting about Brown and Kulik's operational definition of flashbulb memories is that it would clearly seem to "pave the way" for the allowance of forgetting in flashbulb memories. However, at the same time, Brown and Kulik suggested that individuals might remember the circumstances surrounding a surprising and consequential event because of the potential biological significance of the event and the relatedness of the event to survival of the species. Thus, the authors emphasized the importance of remembering the event to a species' very survival and, further, suggested a photographic metaphor to describe the mechanism by which these types of memories might be formed. Even the term used by these authors, *flashbulb memories*, suggests that such memories have rich details that are frozen in time, much like a photograph. Thus, depending on which of Brown and Kulik's conceptualizations one focuses on, he or she may end up with the understanding that flashbulb memories are special and involve little forgetting or the understanding that flashbulb memories may be special, but, perhaps, not so special.

Since Brown and Kulik's study, several investigators have examined individuals' reactions to surprising and consequential events to determine the nature of flashbulb memories. One important debate in the literature has been whether flashbulb memories are different enough from other memories to suggest that they must be formed differently than are other memories. If so, flashbulb memories for important events in people's lives (and the circumstances of hearing about the events) may be impervious to the types of rampant forgetting that individuals experience for more mundane events. To examine the "special mechanism" hypothesis, psychologists have frequently examined individuals' ability to recall the circumstances of hearing about a surprising and consequential public event fairly soon after hearing of the event and then again, after a longer delay. For example, individuals

might be asked a series of questions regarding their memory of hearing about the *Challenger* explosion a day after the news event and then again 6 months or 1 year later. Of interest is whether individuals remember the event and the circumstances of hearing about the news event and whether their memories of the event and circumstances after a delay are similar to those they had initially. It is generally assumed in this type of research that individuals' initial recollections, typically examined within days of the event, are accurate.

When psychologists have examined individuals' memories of public events in this way, it has generally been the case that individuals' memories are not consistent over time and involve a loss or distortion of details. Some recent research suggests that these types of changes are less likely to occur when the initial memories are investigated about a week or so after the event, rather than immediately following the event. In any case, it appears that flashbulb memories are not so special as to warrant the hypothesis that they are formed from a special set of processes. Some psychologists have found, however, that individuals often have a great deal of confidence in their inaccurate memories. Thus, although the individuals' memories for important events may not be accurate, these individuals believe that their memories are accurate.

Another important issue is what factor or factors are necessary to form a flashbulb memory. Psychologists have examined the role of several variables in the formation of flashbulb memories, including rehearsal, emotional state, importance or consequentiality, surprise, novelty, and prior knowledge. Psychologists have also developed and tested fairly complex models of how these variables might interact to produce flashbulb memories. It does seem clear at this point that some of these variables may play a more crucial role in the formation of flashbulb memories than others. Psychologists are continuing to refine their models of flashbulb memory formation.

Although it is now generally agreed upon that flashbulb memories involve forgetting, it is also the case that individuals' memories for some events, and the circumstances of hearing about the events, can be impressive. There have been cases in the literature, for example, of older adults recalling a striking amount of details for crucial historical events that occurred decades earlier. For this reason, psychologists have begun to concentrate on identifying the circumstances leading to such superior memory. It

appears that rehearsal may be critical in the formation of flashbulb memories, as individuals retell important narratives about their life history. It also appears that direct involvement in an event may be an important component in the formation of flashbulb memories. For example, living through an event such as an earthquake or being in close proximity to such an event leaves a much stronger impression than hearing about this event from many miles away. Thus, there is a difference in having "been there" for something like the Loma Prieta earthquake or the 9/11 attacks in New York City. As research on flashbulb memories continues to progress, psychologists will no doubt discover other important factors related to the formation of flashbulb memories and gain an even greater understanding of what factors are responsible for the very impressive memory for circumstances that may occur in some situations.

Karen M. Zabrucky and Lin-Miao L. Agler

See also Emotion and Memory; Episodic Memory; Long-Term Memory; Memory

Further Readings

Brown, R., & Kulik, J. (1977). Flashbulb memories. *Cognition*, 5, 73–99.

Finkenauer, C., Luminet, O., Gisle, L., El-Ahmadi, A., van der Linden, M., & Philippot, P. (1998). Flashbulb memories and the underlying mechanisms of their formation: Toward an emotional-integrative model. *Memory and Cognition*, 26(3), 516–531.

McCloskey, M., Wible, C. G., & Cohen, N. J. (1988). Is there a special flashbulb-memory mechanism? *Journal of Experimental Psychology*, 117(2), 171–181.

Neisser, U., & Harsch, N. (1992). Phantom flashbulbs: False recollections of hearing the news about *Challenger.* In E. Winograd & U. Neisser (Eds.), *Affect and accuracy in recall: Studies of flashbulb memories* (pp. 9–31). New York: Cambridge University Press.

Otani, H., Kusumi, T., Kato, K., Matsuda, K., Kern, R. P., Widner, R., et al. (2005). Remembering a nuclear accident in Japan: Did it trigger flashbulb memories? *Memory*, 13(1), 6–20.

Smith, M. C., Bibi, U., & Sheard, D. E. (2003). Evidence for the differential impact of time and emotion on personal and event memories for September 11, 2001. *Applied Cognitive Psychology*, 17, 1047–1055.

Winningham. R. G., Hyman, I. E., & Dinnel, D. L. (2000). Flashbulb memories? The effects of when the initial memory report was obtained. *Memory*, 8(4), 209–216.

FLUID INTELLIGENCE

Fluid intelligence is the set of cognitive processes that people bring to solving novel tasks and representing, manipulating, and learning new information. Consequently, fluid intelligence is an important construct in educational psychology because it attempts to describe and explain aspects of the individual that influence how, and how well, people solve unfamiliar problems and learn previously unfamiliar material. The history, nature, and current controversies surrounding fluid intelligence are herein reviewed.

History

Early research in intelligence proposed that intelligence was composed of a single, unitary characteristic (known as general intelligence, or g) and a relatively large number of specific abilities. Whereas g was viewed as broad ability having a profound effect on learning, problem solving, and adaptation, specific abilities were viewed as narrow and largely trivial. However, subsequent research differentiated intellectual abilities that were based, in large part, on culturally specific, acquired knowledge (known as crystallized abilities, or g_c) and intellectual abilities that were less dependent on prior knowledge and cultural experiences (known as fluid abilities, or g_f). Although this work was primarily influenced by factor analysis of relationships among cognitive tests, prediction of future learning, experimental studies, and other forms of evidence also supported the crystallized versus fluid distinction. More modern research has identified other abilities in addition to crystallized and fluid abilities (e.g., working memory, quantitative reasoning, visualization), although scholars have not yet agreed on the exact number and nature of these abilities and whether these abilities are independent faculties or subordinate to g. In contrast, there is strong consensus on the distinction between fluid and crystallized intellectual abilities and their substantial roles in human learning and adaptation.

Nature

Contemporary neuroscience defines fluid intelligence as cognitive processing independent of specific content. Fluid intelligence is characterized by the ability to

suppress irrelevant information, sustain cognitive representations, and manage executive processes. Measures of fluid intelligence are strong predictors of cognitively demanding tasks, including learning, education, vocational performance, and social success, particularly when such performance demands new learning or insight rather than reliance on previous knowledge.

Research also suggests strong biological influences on the development of, and individual differences in, fluid intelligence. For example, studies demonstrate that (a) fluid intelligence is more heritable than most other cognitive characteristics; (b) localization of fluid intelligence operations in the prefrontal cortex, anterior cingulate cortex, amygdala, and hippocampus; (c) life-span changes associating neurotransmitter decreases with decrements in fluid intelligence; (d) moderate associations between neural speed of response/conduction and (untimed) measures of fluid intelligence; (e) that unusual exposure to language (e.g., deafness, nonstandard language background) has little effect on the development and performance of fluid intellectual abilities; and (f) fluid abilities have been rising steadily in Western countries for over a century in contrast to relatively stable crystallized abilities (i.e., the Flynn effect). Although there is an association between environmental advantages (e.g., parental education, socioeconomic status) and fluid intelligence, this association may be partly or entirely explained by gene-environment correlations. There is little evidence to suggest that deliberate environmental interventions (e.g., compensatory education programs) substantially influence fluid intelligence, although such programs may have at least short-term effects on crystallized intelligence.

Nearly all major clinical tests of intelligence include measures of fluid and crystallized intelligence. Most notably, tests that historically invoked different models of intellectual processes have recently adopted a hierarchical model in which measures of fluid and crystallized intelligence (and sometimes other abilities) are viewed as subordinate to general intelligence and are combined to produce a composite estimate of g.

Current Controversies

There are a number of unresolved issues regarding fluid intelligence. One such issue is how fluid intelligence relates to other forms of intelligence. Some have argued that general intelligence is a statistical artifact (i.e., there is no genuine all-purpose intellectual ability captured by g) and that fluid intelligence is an independent, distinct intellectual ability (i.e., the multiple intelligences perspective). Although the notion of multiple intelligences is popular among educators, most cognitive scientists hold that abilities are hierarchically arranged, with g as a superordinate construct. Some hold that fluid intelligence is synonymous with g and that g_f is, therefore, the superordinate construct in any hierarchy of intellectual abilities, but most hold that g is distinct from, and superordinate to, g_f.

Another controversy is whether measures of fluid intelligence provide a more "culturally fair" approach to estimating the intelligence of individuals from linguistically, ethnically, or culturally diverse backgrounds. Cross-cultural and cross-linguistic research generally supports this contention, as fluid intelligence tests can typically be used with few changes and still yield reliable and valid results. However, the case for ethnic groups sharing a common language (e.g., English-speaking Black and White Americans) is more controversial. Whereas measures of crystallized intelligence produce substantial differences between groups and are often cited as examples of cultural bias in assessment, the finding that measures of fluid intelligence reflect similar (and in many cases, larger) differences between groups is not as widely recognized. Although at this time, the bulk of evidence is consistent with the conclusion that tests reflect, rather than create, intellectual differences within and between groups, some scientists continue to search for explanations of differences between groups that attribute the differences to nonintellectual factors.

The processes that constitute fluid intelligence are also a matter of debate. For example, some argue that working memory and fluid intelligence are synonymous, whereas others argue that fluid intelligence is a unique characteristic related to, yet distinct from, other cognitive processes. Yet another controversy relates to the methods used to decide competing theories. For many decades, the dominant source of evidence was factor analytic. However, more recent research has invoked brain imaging, functional magnetic resonance, genetic, and other forms of evidence. The generation and integration of evidence from these disparate sources promise to help resolve current controversies, and pose new controversies, regarding the nature and definition of fluid intelligence.

Jeffery P. Braden

See also Crystallized Intelligence; Intelligence and Intellectual Development; Intelligence Quotient (IQ); Intelligence Tests

Further Readings

Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working memory and intelligence: The same or different constructs? *Psychological Bulletin*, *131*(1), 30–60.

Blair, C. (2006). How similar are fluid cognition and general intelligence? A developmental neuroscience perspective on fluid cognition as an aspect of human cognitive ability. *Behavioral and Brain Sciences*, *29*, 109–160.

Conway, A. R. A., Cowan, N., Bunting, M. F., Therriault, D. J, & Minkoff, S. R. B. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence*, *30*, 163–183.

Gray, J. R., Chabris, C. F. & Braver, T. S. (2003). Neural mechanisms of general fluid intelligence. *Nature Neuroscience*, *6*, 316–322.

Johnson, W., & Bouchard, J. T., Jr. (2005). The structure of human intelligence: It is verbal, perceptual, and image rotation (VPR), not fluid and crystallized. *Intelligence*, *33*, 393–416.

Neisser, U. (Ed.). (1998). *The rising curve: Long-term gains in IQ and related measures*. Washington, DC: American Psychological Association.

Neisser, U., Boodoo, G., Bouchard, T. J., Jr., Boykin, A. W., Brody, N., Ceci, S. J., et al. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, *51*, 77–101.

FREQUENCY DISTRIBUTION

It is not only important to understand what descriptive data represent but, if possible, to see it as well. One way to do this is through the use of a *frequency distribution*, a visual representation of a distribution of data.

Table 1 shows 25 scores on a math test for which a frequency distribution will be created.

Table 1   Sample Math Scores

70	77	80	84	90
72	78	80	84	91
72	78	82	85	91
74	79	83	87	93
76	80	84	87	94

The first step in creating a frequency distribution is to define the class interval that will be used. A *class interval* is a range of numbers, and the first step in the creation of a frequency distribution is to define how large each interval will be.

Some guidelines for creating a class interval are as follows:

1. Select a class interval that has a range of 2, 5, 10, 15, or 20 data points.

2. Select a class interval so that 5 to 20 such intervals cover the entire range of data. A convenient way to do this is to compute the range, then divide by a number that represents the number of intervals you want to use (between 10 and 20).

3. Begin listing the class interval with a multiple of that interval.

4. Finally, the largest interval goes at the top of the frequency distribution.

Once class intervals are created, it is time to complete the frequency part of the frequency distribution. This is done simply by counting the number of times a score occurs in the raw data and entering that number in each of the class intervals represented by the count.

In the frequency distribution created earlier, the number of scores that occur between 80 and 84 and are in the 80–84 class interval is 8. So, an 8 goes in the column marked Frequency.

Table 2 shows the frequency distribution following the guidelines listed previously where all 25 scores are represented and can, of course, only appear in one interval.

Table 2   Sample Math Scores–Frequency Distribution

Class Interval	Frequency
90–94	5
85–89	3
80–84	8
75–79	5
70–74	4

Another way to visualize a distribution of scores is through the creation of a histogram such as that depicted in Figure 1.
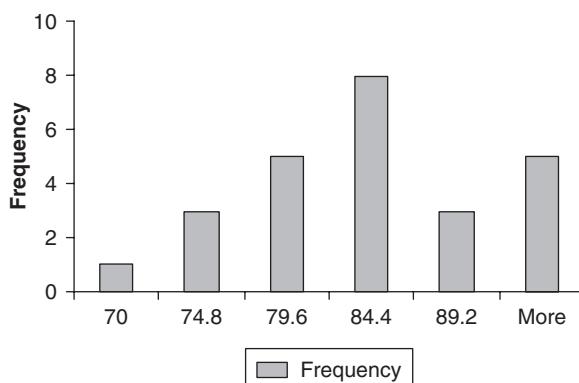
Figure 1   Sample Math Scores—Histogram

Neil J. Salkind

See also Descriptive Statistics; Quantitative Research Methods

Further Readings

Salkind, N. (2003). *Statistics for people who (think they) hate statistics.* Thousand Oaks, CA: Sage.

Warner, R. (2007). *Applied statistics.* Thousand Oaks, CA: Sage.

FRIENDSHIP

Making friends, keeping friends, and being a friend are considered important developmental tasks from early childhood to the adolescent years. Friendships set the stage for children's development of numerous competences, including communication and cognitive skills, as well as emotion regulation and emotion understanding. Friendships also allow children to measure themselves against others, to develop a sense of self-concept, and to acquire the social skills they will use throughout their adult lives. This entry discusses the role that friendships play in school adaptation, in being accepted, and in childhood development. This entry then addresses deleterious effects and other implications of friendships.

Friendships and School Adaptation

Friendships appear to play a crucial role in children's adjustment and adaptation to school. One way

that friendships may contribute to children's school performance is by directly stimulating cognitive growth and learning, thus setting the stage for later intellectual performance. Consistent with this proposal, empirical evidence suggests that children demonstrate greater problem-solving ability, task mastery, and creativity when interacting with friends than when working alone and that these skills transfer to other situations. When children collaborate with friends, they are more efficient and productive problem-solvers across a variety of tasks, including creative and oral tasks, as well as more academic tasks, such as scientific reasoning problems or writing assignments. Further evidence links the quality of children's friendships to children's academic performance. Such associations appear as early as the preschool years, with the quality of preschool children's friendships predicting academic performance in elementary school.

A second way that children's friendships may influence adjustment to school is by shaping children's attitudes and motivation toward schooling. For example, research by Carollee Howes indicates that children who moved from one day care setting to another, accompanied by friends, demonstrated higher levels of social competence than children making a similar transition without a friend. Brian Vaughn and his colleagues report a similar finding for preschool children advancing from one Head Start classroom to another. Gary Ladd and his colleagues document the positive value of entering kindergarten in the company of friends on children's adjustment. Particularly noteworthy is their finding that children who moved into kindergarten with mutual friends were better adjusted to school than children who made the transition with acquaintances who were not friends. The transition process from elementary to middle school also appears to proceed more smoothly for children who have friends. Specifically, Kathryn Wentzel found that students who had a friend upon entering middle school displayed better academic and social adjustment at the end of their first year of middle school than students without a friend. In addition to helping children cope with the normative stress that accompanies the transition to a new school, evidence suggests that friendships may help children cope with nonnormative stresses such as becoming the victim of a school bully. In contrast, children without friends not only report being lonely and feeling depressed but also exhibit inappropriate classroom behavior.

Friends, Acquaintances, and Peer Acceptance

The impact that friendships have on children's school adjustment may result from the uniqueness of the friendship relationship compared to children's other social relationships. A great deal of research has been devoted to identifying differences between children's interactions with acquaintances versus friends. Evidence from this body of work suggests that friends feel a sense of responsibility for one another's needs, that they assist each other in meeting those needs, and that they expect no repayment of the assistance they provide to their friend. In contrast, acquaintances tend to feel less responsibility for one another and are more likely to distinguish the circumstances under which they will provide support to each other. In a comprehensive meta-analysis of studies comparing friends and nonfriends, Andrew Newcomb and Catherine Bagwell concluded that friends engage in more frequent positive interactions, including talking, cooperation, and positive affect, than do peers not identified as friends. Friends are also more similar behaviorally to one another, more egalitarian, more loyal to one another, and less likely to assert dominance over one another than are children who are simply acquaintances. These characteristics are presumably a consequence of friends' greater proximity, heightened mutual interest, and intense concern for one another, and they point to the unique affiliative bond that friends share.

A particularly noteworthy behavioral domain in which friends differ from acquaintances is conflict management. Although friends engage in conflict at rates similar to those of nonfriends, friends are distinguished from nonfriends in their conflict resolution efforts. For example, Willard Hartup and his colleagues found that, compared with conflicts between children who were not friends, conflicts between (4-year-old) friends were less heated, were more likely to be resolved by both partners disengaging from the conflict, were more likely to end in compromise, and were more likely to be followed by continued interaction. The authors suggest that because friendships are founded on emotional commitments, children have a lot at stake in their friendships, and failing to resolve conflict puts the friendship at risk. Consequently, children appear to be particularly motivated to reduce conflict with their friends in order to maintain the relationship. This evidence joins with other research documenting differences in the behavioral

qualities of friendship and children's interactions with acquaintances to suggest that friendships offer children unique experiences, provide distinct developmental resources, and serve different functions than do children's other relationships.

Friendships have also been distinguished from children's social status, or level of acceptance, in the larger peer group. Acceptance refers to being generally well-liked by a group of peers, rather than participating in a specific, close dyadic relationship with one other individual. Research with school-age children and adolescents suggests that acceptance and friendship are unique but related domains of children's peer relationships. For example, in a study examining both friendship status and peer acceptance in early childhood, Ladd found that children who formed new friendships in their kindergarten classroom experienced improvements in school performance over the course of the school year. However, even after taking into account children's friendship status, peer rejection predicted less favorable attitudes toward school, greater school avoidance, and lower levels of academic performance at the end of the school year. More recently, Ladd and Wendy Troop-Gordon found that children who were friendless in first grade had more internalizing adjustment problems and reported being more lonely at school in fourth grade than children with friends. Moreover, the connection between being friendless and children's internalizing problems held true even after taking into account children's chronic peer rejection. Thus, it appears that being accepted by peers and having friends may provide children with different developmental opportunities and may have different connections to children's adjustment in school.

At the same time, however, popularity with peers appears to overlap with the development of friendships. In their summary of research on connections between children's acceptance among peers and participation in reciprocated friendships, Willard Hartup and Nan Stevens concluded that friendship is a positive correlate of social competence for preschool children. Consistent with this view, Brian Vaughn reported that the number of reciprocated friendships enjoyed by preschool children was a positive and significant correlate of overall social competence scores. One way that peer acceptance appears to influence friendships is by determining the amount of choice that children have in friends. Those children who are more popular tend to have more friends, engage in

more cooperative activities, and display more social conversation than children who are socially rejected or neglected. At the same time, however, being well-liked by peers does not guarantee that a child will have a specific friendship, nor does being rejected by one's peer group preclude a child from having a friend. A recent study of second and third graders found that 39% of children rejected by their peer group had at least one mutual friend in that group and that 31% of popular children did not have a reciprocated friendship.

Although there is some question as to when children form friendships, evidence suggests that children begin to discriminate among peer partners and form preferences for particular playmates within their peer groups as early as toddlerhood. Children as young as 30 to 73 months of age (nominal "preschool" ages) have been confirmed to show preferences for specific peers in their day care and preschool groups. These early friendship preferences manifest themselves by the maintenance of close proximity and frequency of interaction between two children. By 36 to 48 months, over half of all children have reciprocated friendships, and many of these friendships are stable over time. Preschool friendship preferences can be reliably assessed using a sociometric choice task, a procedure that asks children to identify their most preferred playmates. Evidence suggests that these early friendships may be differentiated from other peer relationships by the amount of time children spend in reciprocal and complementary interaction.

Friendships and Childhood Development

Age plays a major role in defining the specific behavioral qualities that distinguish children's friendships from other peer relationships and in shaping expectations for interpersonal behavior within friendships. Based on well-defined physiological, cognitive, emotional, and social changes, development during childhood is divided into three major periods: early childhood (3–7), middle childhood (8–12), and adolescence (13–18). Throughout these periods of development, friendships are characterized by reciprocities, mutual liking/attraction, and by affection and having fun together. However, there are marked qualitative differences in children's friendships during different age periods, and the behavioral processes involved in the formation and maintenance of friendships change as children develop. Consequently, friendship interaction during any given developmental period has a characteristic structure and content that give it a distinct signature or theme. The themes of each period and their developmental progression are normative in that they reflect children's attempts to adapt to the demands of their social-ecological niche.

Early Childhood

During early childhood, friendships are defined by the theme of playmate. Young children place a premium on a child's potential as a playmate, both in their descriptions of actual friendships and in their beliefs about friendships in general. Consequently, in this developmental period, friendships may be identified by children's maintenance of close proximity, engagement in mutually interesting activities, and time spent together. The level of enjoyment and satisfaction children experience with playmates depends almost completely on the level of coordination they achieve. Coordination refers to the extent to which two children are able to fit together their separate actions into jointly produced discussions or activities. Coordination is easier to achieve when children share similar interests and behavior patterns. For this reason, similarity is a key factor in young children's friendship formation. However, it is not the specific nature of their similarities that interest young children, but rather the mere presence of commonalities. In their effort to establish commonalities, a good deal of young children's friendship conversation involves social comparison processes. Once friendships are formed on the basis of commonalities, efforts to maintain those relationships are motivated by the desire to achieve coordination of play. Thus, the interactive emphasis of young friends centers on play negotiation, emotion regulation, and conflict management.

Middle Childhood

As children transition from early childhood to middle childhood, their conceptions of friendship change from a self-centered view of what a friend can do for them as an exciting playmate to a more balanced perception of friendship, involving reciprocal and mutually satisfying interactive exchanges. In addition, children move developmentally from an emphasis on friendship being a momentary act between individuals to a focus

on the enduring affective nature of the relationship. Consequently, during middle childhood, children's friendships have been characterized by the theme of companionship. During this period, friendships become more stable and are more often reciprocated. Also by middle childhood, as a function of growing interpersonal awareness, the basis of friendships centers more on shared norms and personal qualities, as opposed to common interests. The behavioral processes that are salient to friendship also change during middle childhood. For example, fantasy play among friends appears to decline dramatically from preschool to middle childhood, whereas animated conversation, games, and contests become more common. In addition, during middle childhood, the content of conversations between friends begins to focus more on sharing intimate information about the self, and children come to expect that friendship creates an obligation for mutual aid or support when needed or requested by a friend.

Adolescence

In adolescence, friendships depend on and manifest themselves in intimate, dyadic exchanges that feature openness, honesty, and affection. Evidence suggests that as adolescents become preoccupied with concerns about developing autonomy and self-governance, they begin to rely more on friends for intimacy and support. Accordingly, the theme of friendships during this developmental period is intimacy. Friendships serve the function of helping adolescents to achieve individual identity and self-understanding. This role is evident in the expectations that adolescents espouse about friendships, which focus on similarities in interests and personalities, and processes of intimacy and validation. Likewise, the interactive emphasis of adolescents' friendships center on self-disclosure, positive gossip, and shared intimacy. During adolescence, friendships become the primary social relationships in which young people share their confidences. Emotional support and autonomy are also seen as essential elements of friendship among adolescents. However, compared to previous ages, adolescents are less preoccupied with what benefits they derive from their friendships and are more conscientious of their friends' needs.

Negative Effects

Although it is clear that friendships typically feature reciprocity, loyalty, shared positive affect, and

companionship at all ages, it is also important to note that friendships can serve as contexts for less-amicable social processes. As previously indicated, conflict is prevalent in interactions between friends, and although friends typically manage conflict in more optimal ways than do children who are not friends, the fact remains that excessive interpersonal conflict is predictive of adjustment problems. Friendships can also serve as the setting for harmful interpersonal processes, such as coercion, jealousy, and betrayal. Moreover, children may form friendships with children who demonstrate antisocial characteristics. Such friends, in turn, can pressure children to engage in behavior they might not otherwise consider. In addition, some friendships can magnify the child's own antisocial tendencies, or they may lead to bully-victim relationships. For these reasons it is important for parents and teachers to consider the quality of children's friendships and to give attention to whom children form friendships with.

Implications

Research clearly indicates that most children have at least one friend. Moreover, empirical evidence points to the significant role that friends play in children's adjustment to school and academic performance. Friendships can be distinguished from other social relationships in which children are involved. Friends engage in interactions that are qualitatively different from those of children who are not friends. Children who are friends communicate clearly, self-disclose more often, and resolve conflicts in ways that enhance the likelihood that their relationship will continue. Friendships also have been distinguished from children's social acceptance in their peer group. A significant portion of children who are well-liked by their peer group do not have a specific friend, and a substantial number of children who are rejected by peers do have a friend. Characteristics of children's friendships change with age. As children grow older, their thinking about friendship progresses from the concrete to the abstract, and this change is manifested in their interactions with friends. Specifically, over time, children's friendships become more stable, demonstrate more reciprocal altruism, and consist of more shared intimate personal knowledge. Recent research also indicates that friendships can have deleterious as well as constructive effects on children's adjustment and can involve both positive and supportive

interactions as well as conflictual and stressful interactions. Moreover, who a child chooses as friends appears to be as important as whether a child has friends or not. The implications of this information concerning children's friendships are complex and important for both educational policy and classroom practices.

Eric W. Lindsey and Heather Stopp

See also Emotional Development; Gender Differences; Peer-Assisted Learning; Peer Influences; School Readiness; Social Development

Further Readings

Hartup, W. W. (1996). Cooperation, close relationships, and cognitive development. In W. M. Bukowski, A. F. Newcomb, & W. W. Hartup (Eds.), *The company they keep: Friendship in childhood and adolescence* (pp. 213–217). Cambridge, UK: Cambridge University Press.

Hartup, W. W., & Stevens, N. (1997). Friendship and adaptation in the life course. *Psychological Bulletin*, *121*, 355–370.

Howes, C. (1988). Peer interaction of young children. *Monographs of the Society for Research in Child Development*, *53*(1, Serial No. 217).

Ladd, G. W. (1990). Having friends, keeping friends, making friends, and being liked by peers in the classroom: Predictors of children's early school adjustment? *Child Development*, *61*, 1081–1100.

Ladd, G. W., Kochenderfer, B. J., & Coleman, C. C. (1996). Friendship quality as a predictor of young children's early school adjustment. *Child Development*, *67*, 1103–1118.

Ladd, G. W., & Troop-Gordan, W. (2003). The role of chronic peer difficulties in the development of children's psychological adjustment problems. *Child Development*, *74*, 1344–1367.

Newcomb, A. F., & Bagwell, C. L. (1995). Children's friendship relations: A meta-analytic review. *Psychological Bulletin*, *117*, 306–347.

Vaughn, B. E., Azria, M. R., Krzysk, L., Caya, L., Bost, K. K., Newell, W., et al. (2000). Friendship and social competence in a sample of preschool children attending Head Start. *Developmental Psychology*, *36*, 326–338.

Vaughn, B. E., Colvin, T. N., Azria, M. R., Caya, L., & Krzysk, L. (2001). Dyadic analysis of friendship in a sample of preschool-age children attending Head Start: Correspondence between measures and implications for social competence. *Child Development*, *72*, 862–878.

Wentzel, K. R., Barry, C. M., & Caldwell, K. A. (2004). Friendships in middle school: Influences on motivation and school adjustment. *Journal of Educational Psychology*, *96*, 195–203.

G

The young do not know enough to be prudent, and therefore they attempt the impossible—and achieve it, generation after generation.

—Pearl S. Buck

GANGS

The topic of gangs has generated extensive literature and social action. Questions over both definitions and solutions typify gang research, policy, and intervention. Law enforcement efforts at containing gangs date back to colonial America. In turn, modern empirical research into gangs began in the late 1920s. Despite the longevity of gangs as a social problem, there is still ongoing debate concerning the composition of gangs, their behavioral dynamics, and the most effective gang reduction strategies.

This entry discusses the definitions of gangs, the factors that cause them, gang typology, and the dynamics of gang membership. This discussion is followed by an examination of the various types of gangs. Finally, differing anti-gang strategies and problem-solving approaches are explored.

Differing Definitions

Probably the biggest challenge to researchers, anti-gang practitioners, and policy makers alike is that there is no standard or widely accepted definition of the word *gang*. Because there are so many competing definitions, any discussion or collaboration between different groups suffers; there is no common vocabulary. Communication is hindered by muddled comparisons and lack of terminological clarity. This confusion in definitions and concepts contributes to greater problems in trying to understand the prolific growth of gangs from the 1990s onward. Furthermore, this confusion contributes to the difficulty building consensus around the most effective anti-gang strategies. This definitional problem raises the question of which comes first: Does gang membership push youth into criminal activity, or does criminal activity push youth into gang membership? Because the concepts underlying membership in gangs are similar to the concepts underlying membership in other social organizations, the result is contradictory definitions and conflicting problem-solving approaches.

Every group involved in solving the gang "problem" agrees that there is a need for a universally accepted definition that would serve as a basis for action, policy, and research. However, this is where agreement ends. Depending on the perspective, the definition of gangs and gang membership differs.

Definitions derived from criminal justice and law enforcement diverge from those utilized in the social sciences and comprehensive academic approach. Although there is some strong overlap between these definitions, each is distinct. Criminal justice definitions focus on violations of the law. Academic definitions focus on the structure and motivation behind gang behavior. Additionally, there is frequently little agreement

on definitions at the federal, state, and local levels of government; laws and policies developed at different levels conflict.

Criminal Justice and Law Enforcement

According to criminal justice professionals, a *gang* is any group that gathers on an ongoing basis to engage in antisocial or criminal activities. Gang members identify with one another based on geographical location, clothing colors, symbols, and names. They communicate their gang affiliation through hand signs and graffiti. Law-breaking activities enhance the gang's credibility, create fear in the community, and may provide an ongoing source of income for the gang and its members. The magic number of individuals necessary to constitute a gang is universally acknowledged to be three. Nevertheless, rigid definitions such as these frequently omit information needed to understand the phenomenon of gang membership and guide both policy and research.

Academics and Social Sciences

Academic definitions are aimed at understanding and communicating information regarding the structure and motivation of gangs. Gangs are viewed as loosely organized groups of individuals who assemble together for social and criminal purposes. Membership in a gang fulfills complex and interrelated psychological needs. Gangs proliferate as a result of social conditions and cultural pressures.

Comprehensive Definition

Gangs can best be defined as any type of ongoing group or organization of three or more individuals, usually adolescents and young adults, who interact frequently with each other and are recurrently and deliberately involved in violence and criminal activity, individually and collectively. This criminal activity includes fights, intimidation, and threats. Gangs may be formal or informal in nature, but all generally form an allegiance for a common purpose. Most notably they share a collective identity that is usually expressed through a gang name as well as identifying signs or symbols. Each gang typically identifies with, and claims control over, certain "turf" or territory in a community—which is not limited to geography but can mean other individuals, objects, or businesses.

Gang membership generally forms along ethnic and socioeconomic boundaries. Long associated with urban settings, gangs have migrated to suburban and rural settings as well. Gangs are now found in all types of communities and socioeconomic strata.

Gangs are loosely controlled by a leader or leaders who direct the gang for their own benefit as well as for the benefit of their followers. Usually, a gang's leaders are characterized by their desire to maintain a strong or "bad" reputation, their need for respect, and their drive to respond to actual or perceived disrespect with acts of retaliation or revenge. A gang's reputation is based on both the actions and influences of its individual members as well as the gang as a collective.

Much debate focuses on the extent of gang organization, structure, and cohesion. One position, held by many in the criminal justice system, maintains that gangs are extremely cohesive and organized with rules, rituals, and stratified social roles. There are dire warnings about the increased migration of street gangs into legitimate business and even politics. The opposite position, predominant among social scientists and gang interventionists, holds that gangs are informal, constantly shifting groups that lack any authentic organization. However, research and practice reveal that both cohesiveness and structure vary with the gang and the setting. Most street gangs are loosely organized, with several members who serve in leadership roles commonly referred to as "shot callers." Membership fluctuates and gang members possess varying degrees of commitment and loyalty to the gang. Gang cohesiveness and feelings of commitment are always highest when the gang is challenged by other gangs or by outsiders—in particular, law enforcement. Some gangs are extremely organized and cohesive, but most remain transient, flexible, and fragmented. Most importantly, there is almost never a central gang authority coordinating gang structure or movement.

Changes in Gangs

Gangs are no longer an exclusively urban phenomenon. Although they remain firmly anchored in urban settings, their presence has spread to rural environments, and many gangs now have transnational affiliation. Additionally, gangs have "aged." No longer restricted to juveniles, their upper age limits for membership have increased, with individuals remaining in gangs for longer time periods. Moreover, gangs are growing more intergenerational. Many possess third and fourth

generation members, the great-grandchildren of "original" gang members. In another shift in exclusivity, gang membership is now no longer limited to males. Additionally, the division between street and prison gangs continues to weaken, with growing integration between formerly separate entities.

Types of Gangs

Deborah Prothrow-Stith offered a typology that describes three major gang types:

1. *Scavenger gangs* are best described as underachiever gangs. These groups are loosely organized, with activities that are largely impulsive and unplanned. Their leadership may shift on a frequent basis. In gang culture, scavenger gangs occupy the lowest status, typified by unsuccessful criminal efforts that are derided by other gangs. Alternatively, these groups can be viewed as "gangs in the making." In fact, many evolve into the more structured territorial gangs.

2. *Territorial gangs* are organized around territory or turf. These are the stereotypical gangs associated with Los Angeles and other urban settings. Their membership is usually well organized. These gangs practice complex, formal initiation rites for new members as well as other ceremonies, rituals, and customs that differentiate members from outsiders. Gang membership is a source of pride, status, and power. The major activity of these gangs is fighting—the crossing of territory or turf lines is used as an ongoing excuse to brawl. These gangs may engage in drug dealing, but only as a secondary activity. Their primary activities remain social and criminal. The drug dealing that occurs is usually a matter of economic survival, not a result of a complex marketing or distribution cartel.

3. *Corporate gangs* are highly structured, criminal entities organized around drug trafficking and profit-making. Their members fulfill the roles and responsibilities of traditional gangsters, sharing characteristics with organized crime more than with inchoate street gangs. These gangs enforce strict codes of secrecy and discipline, with severe consequences for transgression. Their membership is highly intelligent and their leadership, akin to any corporate entity, exhibits sophisticated strategic planning as well as financial and personnel management.

Description and Causes

Gang Member Profile

The mystique of gangs has generated myths concerning the "typical" gang member. In truth, the typical gang member may be uneducated but above average in intelligence. Most gang members exhibit sophisticated street survival skills, and many are accomplished in the art of manipulation. Additionally, they follow a strict code of conduct as defined by the gang. Each gang member is concerned with proving their worth, being recognized for their work, and establishing a reputation for being "bad." They are loyal to all other gang members, their loyalty bordering on a family-like affinity, and they are very protective of their turf. This does not minimize the antisocial, destructive core of their behavior but rather places it within context.

Etiology

Gangs find their genesis in the sociological and economical conditions that foster and reinforce individual and psychological vulnerabilities. The three predominant, interrelated factors promoting the existence of gangs are family dysfunction, community violence, and extreme poverty. Children growing up in gang-saturated and economically deprived urban "hot spots" often perceive few alternatives to the gang lifestyles. However, as gangs have spread to rural, nonurban environments, the common denominator in their continued existence and further proliferation is family disintegration. The gang replaces the family as the new source of attachment, security, and stability. By fulfilling these basic human needs, gangs solidify their control over individual gang members.

Characteristically, gang members come from uninvolved or dysfunctional families. Their family relationships are unstable, and many have parents or siblings who encounter ongoing problems with substance abuse. Immediate and extended family members are often involved in gang or criminal activities. Many youth who join gangs have been physically, emotionally, or sexually abused or have suffered extreme child neglect. Gang membership appears as an attractive alternative to the child welfare system and likely placement in foster care. Equally dangerous enticements arise when parents, older siblings, or other relatives have been, or continue to be, gang members. These family members embody role models for young individuals to emulate.

All individuals possess needs for feelings of self-worth, identity, acceptance, recognition, purpose, and

security. Gangs often supply what traditional systems have failed to provide. Youth who are frequently poor, uneducated, and adrift experience the gang "surrogate family" as a source of security and identity in the absence of other support systems. The gang offers acceptance and structure, which—when contrasted with the instability of family life—presents an alluring substitute. Psychological attachment paired with social and environmental factors provides strength to the dynamics of gang membership.

Gang youth believe, often defensively, that they will never succeed in conventional society. Disenfranchised, they lack even the most basic employment skills, such as maintaining a proper appearance or demeanor. However, more fundamentally, gang members feel disconnected from mainstream social and cultural values. They do not perceive the future as filled with opportunities. Additionally, individuals believe that illegal activities provide the only means by which to earn money easily without working.

Gang membership offers a sense of empowerment and status. Members who may have dropped out of school and are unemployed obtain a sense of purpose, identity, and occasionally even ad hoc leadership training. For many individuals, this constitutes a heady mixture of excitement, power, and success. This is integrated with a need for acceptance that may also be fueled by intimidation from others. Many join gangs out of the need to belong or feel accepted by others; young women may join because they have boyfriends in a gang. In more severe circumstances, an individual may join a gang as a response to violent peer pressure from within the gang or as a form of protection from violent threats from rival gang members.

For many in gangs, there is a sense of fatalism that masks their feelings of aimlessness and lack of control. This fatalism often translates to a preoccupation with death and how one dies. Many gang members report dreaming of their deaths, how they wish their funerals be conducted, and their despair at not living to see their children grow up. This morbid preoccupation is a motivating factor in the contradiction of having children while, at the same time, participating in the acts of violence that are hallmarks of gang membership.

Dynamics of Gang Membership

The evolutionary process that brings a child or youth from innocence to gang membership involves a gradual yet powerful set of forces. Gangs work on recruiting vulnerable youth and monitoring their progress. These recruits are slowly allowed to associate with the gang. When gang leadership determines that the youth is ready for full membership, elaborate initiation rituals are imposed to test loyalty, extinguish individuality, and promote service to the gang. The most common initiation ritual of being "jumped in" includes being beaten, humiliated, and, if the gang member is female, raped by multiple gang members. Occasionally, a desirable member is "courted in" and proffered regular gang membership without any initiation ritual.

Gang Membership Typology

Gang membership can best be divided into five major categories:

1. *Hard-core* gang members compose approximately 10% of the gang population. Hard-core gangsters usually have been gang members for the longest time periods and are influential in directing gang activities. They are frequently in and out of jail, unemployed, and using and selling drugs.

2. *Regular* gang members are the younger gang population, ranging in age from 14 to 20 years. Having been initiated or "jumped" into the gang, they back up hard-core gang members in word and action. Those with juvenile status often act as lead participants in gang activities because it is believed that the juvenile justice system will be less punitive. In this way, their commission of crime protects older gang members. In many cases, they consider themselves to be veritable hard-core gangsters "in training."

3. *Associate* gang members can be characterized as ambivalent concerning their gang status. They believe there is no way to exit the gang. These individuals exhibit a certain degree of moral awareness as to the cost of their involvement. However, due to a combination of intimidation, fear, and loyalty, they remain unable to renounce membership and unresponsive to intervention efforts.

4. *Wannabe* gang members vary in age, but generally range from 10 to 13 years. Anecdotal evidence suggests even younger children are increasingly aware of and attracted to gangs. These individuals are not formal gang members but act as if they are in the gang. They dress in gang attire, spend time with gang members, and tag (i.e., write graffiti associated with the gang on surfaces visible to the public).

5. *Potentials* are individuals who live in or close to areas populated by active gangs. These youth may have a family member who already belongs to a gang. However, individuals in this category perceive that there are alternatives, such as school, which can be accessed in order to avoid gang membership altogether.

Consequences of Gang Membership

The price of gang membership is high, and consequences are extreme. From the onset of membership, violence is a hallmark of life in the gang. Initiation rituals involve several forms of physical force, including being "jumped in." This involves a group of gang members physically beating the youth. In addition to this beating, there is typically a requirement that the individual commit serious criminal and violent acts to prove his or her loyalty. Once a youth is initiated, gang membership invariably guarantees a criminal record along with the physical risks associated with ongoing violent activities. It is also important to remember that labeling an individual a gang member is a stigmatizing event. In general, mainstream culture is extremely judgmental, and many believe that once an individual belongs to a gang, that membership is retained for life.

Usually, gangs depend on their youngest regular members to carry out the most serious offenses in the mistaken belief that juveniles receive more lenient treatment when found guilty of a crime. Instead, following the passage of anti-gang laws in several states, this is no longer the case. Youth are now legally criminalized at progressively younger ages. Correspondingly, youth are now socially assigned a career of gang membership at earlier ages. This phenomenon is reflected in both gang recruitment efforts and gang prevention programs, both of which are aimed at younger-aged children.

Violence

There is a great deal of disagreement in the theoretical literature about the association between gangs and violence. However, although reports on the day-to-day existence of gangs reveal that violence is not perpetual, statistics point to a high association between gangs and ongoing violent crime.

Dynamics of Gang Violence

The potential for gang violence occurs most frequently during gatherings in which gang members enter the turf of a rival gang. It is an oversimplification to assume that this occurs only on certain streets or in certain neighborhoods under the control of a gang. In reality, such "turf violations" may also occur in regional centers for sports, recreation (parks), entertainment (movie theaters and concert venues), and schools. Any bus stop, street, or freeway is also "fair game" in terms of gang violence if a gang crosses turf lines.

Potential gang violence may also erupt over the issue of "disrespect," which assumes many forms. One common form of disrespect involves tagging and occurs when one gang crosses out the graffiti of another gang. Menacing looks or stares, "throwing" gang signs, hand signals, and other nonverbal confrontations are also invitations to violence. Verbal precursors include talking trash, ragging on, bragging, and calling out gang names as well as taunts or threats. Other forms of intimidation, such as stepping on toes, bumping, or other "accidental" or intentional physical encounters can spark an incident. Though childish in quality, the smallest incident often can be perceived as disrespectful and constitutes a virtual call to arms. Indeed, the randomness of gang violence is exacerbated by the impulsivity with which gang members determine what actions are disrespectful as well as what to overlook.

Guns

Gang involvement or membership is strongly correlated with gun possession and use. This is due to combined factors: the growing availability of weapons, intimidation by other gangs, the need for self-protection, and the quest for status and identity. However, the power and protection provided by gun possession also pose an added threat: Gang members are at greater risk for being harmed or killed by someone with a weapon.

Increasingly, gangs acquire and use automatic and semiautomatic weapons. These are employed in drive-by shootings, as gang members in a motor vehicle "drive by" and shoot at rival gang members. The use of such lethal weapons has altered gang activity, rendering it more deadly. Painful and unanticipated losses have resulted from this lethality, particularly when gang members miss their intended targets and instead hit uninvolved community members, including children.

Drugs

Both law enforcement and social researchers have sharpened their focus on drugs as a mainstay of gang

activity, revenue, and power. There are few data on the actual relationship between drugs and gangs. Long-term involvement in drug sales and distribution entails secrecy and cohesion not traditionally associated with gang organization and membership. Nevertheless, gangs may be involved in ongoing, albeit erratic, drug usage and dealing. There is considerable disagreement about whether or not drug dealing is separate or part of an ongoing gang business enterprise. Many law enforcement professionals insist that gangs are businesses with elaborate systems for managing cash flow, laundering money, collecting "taxes," and conducting drug wars. In reality, although some gang homicides arise from disputes over drugs, most violent incidents involving gang members evolve from fights over turf, status, and revenge. Also, most gangs are not akin to the traditional tightly organized, drug distribution cartels. Subgroups within gangs deal drugs, with varying degrees of organization and success. Although the extent of involvement varies, it is reasonable to conclude that drugs are an ongoing and growing component of gang life.

Problem-Solving Approaches

Four major approaches underlie most anti-gang strategy:

1. Suppression
2. Prevention
3. Intervention
4. Comprehensive approach

Overall, the general solution to gang membership and its associated criminal activity involves the use of multifaceted and comprehensive anti-gang programs. These programs include alternatives for at-risk youth, strong law enforcement, and committed community leaders all collaborating to combat the gang culture and offer youth options. However, before the comprehensive approach is examined further, each of the other three anti-gang strategies and the impact each has on reducing gang membership and activity are discussed.

Suppression

The goal of suppression is to reduce gang activity. Based on traditional criminal justice methods, its practitioners, usually comprising law enforcement

officers, probation officers, and prosecutors, use arrest and incarceration to control gang violence and suppress criminal activity. Suppression practitioners enact gang injunctions to keep identified gang members from assembling in public. They also utilize vertical prosecution along with sentencing enhancements for gang involvement, gang membership, or both.

Prevention

Prevention is designed to identify high-risk children and youth who have *not yet* joined a gang and to preclude their joining a gang or participating in gang activity. Primary prevention is aimed at broad populations, for example, children from birth to 5 years old who are growing up in a gang-saturated neighborhood. On the other hand, tertiary prevention programs specifically focus on youth who are gang potentials or gang associates.

Currently, prevention efforts concentrate on middle school youth, with some additional programs developed for elementary school children. Strategies include offering mentoring and academic support as well as activities to initiate and enhance parental involvement. Additionally, sports and recreation as alternatives have been developed in community settings.

When a setting is *not* in crisis, prevention programs command the greatest attention. Once gang crime develops, suppression assumes the predominant role in reducing gang activity—with the sentiment that prevention can be practiced once neighborhood safety is reestablished.

Intervention

Gang intervention encompasses a broad range of actions and strategies designed to constructively influence and divert either individual gang members or one or more gangs in a locale or setting. It is aimed at individuals or neighborhoods that have progressed past the point of being labeled "at-risk."

Gang intervention strategies are limitless in design and focus; some strategies involve individuals who are currently active gang members whereas other interventions focus on individuals reentering society after release from incarceration. Individual intervention includes tutoring, job training, tattoo removal, counseling, assistance with substance abuse, legal aid, and using both community- and faith-based groups to convince gang members to leave the gang.

Group gang intervention brings rival gangs to the negotiation table by employing former gang members as street "interventionists." These street workers engage in dialogue and relationship building that may entail dispute mediation or conflict transformation through the construction of gang truces or peace treaties. Also, when death occurs due to gang violence, gang interventionists often raise money to help families with funeral expenses and concurrently work to curb retaliation and further violence.

Comprehensive Approach

The current cynosure of anti-gang strategy is a combination of suppression, prevention, and intervention known as the *comprehensive approach.* This strategy entails a dynamic collaboration between criminal justice and social service professionals, academicians along with mental health facilities, schools, and community- and faith-based organizations. Equal weight is assigned to community mobilization, workplace development, and law enforcement. Extensive services to all sectors are offered. This approach is potentially effective but also very costly.

Future Directions

The topic of gangs remains dynamic and changing. Both gangs and the research and policy directed toward them are constantly evolving. It is clear that the existence of gangs continues to be traced to a combination of psychological, cultural, and social factors. In recent developments, however, their activities involve increasing levels of violence, diverse groups, and new, nonurban settings. What is most striking about gangs is that they continue to exist and remain resistant to traditional anti-gang strategies. Challenges endure in terms of future gang policy, interdiction, and research. The questions of what a gang is, how membership and organization are structured, and which problem-solving approaches are the most effective are still open to debate.

Jorja Leap

See also Child Abuse; Identity Development; Peer Influences; Poverty; School Violence and Disruption

Further Readings

Egley, A., Jr., Maxon, C., Miller, J., & Klein, M. (Eds.). (2006). *The modern gang reader* (3rd ed.). Los Angeles: Roxbury.

Fremon, C. (2004). *G-Dog and the homeboys: Father Greg Boyle and the gangs of East Los Angeles.* Albuquerque: University of New Mexico Press.

Huff, C. R. (Ed.). (2002). *Gangs in America.* Thousand Oaks, CA: Sage.

Klein, M., & Maxson, C. (2006). *Street gang patterns and policies.* New York: Oxford University Press.

McCall, N. (1994). *Makes me wanna holler: A young black man in America.* New York: Vintage Books.

Prothrow-Stith, D. (with Weissman, M.). (1991). *Deadly consequences.* New York: HarperCollins.

Simpson, C. (2005). *Inside the Crips.* New York: St. Martin's Press.

Spergel, I. A. (1995). *The youth gang problem: A community approach.* New York: Oxford University Press.

GENDER

The word *gender* originally existed as a grammatical category stemming from the Latin word *genus* and can be translated as "kind" or "sort." Genders were known as classes of nouns reflected in behavior of associated words. In some languages, gender is central and adjectives and verbs show gender agreement; in other languages, adverbs, numerals, and other parts of speech agree. In some languages, gender is absent. Stemming from linguistics in which gender referred to feminine and masculine forms within language, gender conveyed strong associations about the role of society in distinguishing words coded as male and female.

During the 1960s, feminist scholars extended the cultural constructions of gendered language to adopt the concept of gender to distinguish cultural characteristics associated with masculinity and femininity from biological features distinct to males and females, such as male and female chromosomes, hormones, and internal and external reproductive organs. The use of the word *gender* rather than *sex* was used to repudiate beliefs that biology determined male and female behavior, including stereotypical beliefs about female inferiority. Gender was distinguished as the cultural part of being a man or a woman, and being masculine or feminine was no longer connected with biological sex but described as a culturally variable characteristic.

Now, in current speech, *gender* has almost completely replaced *sex* except when referring to sexuality. When asking if a person is a man or a woman, it is customary to ask for "gender" rather than "sex." In fact, many psychologists use the term *sex*

differences and *gender differences* interchangeably. However, it is important to clearly distinguish the terms *sex* and *gender* when discussing social psychology. Many influential researchers in the field use *sex* to refer to the binary categories of male and female and *gender* to refer to the attributes associated with the two sexes, including gender roles and gender stereotypes. It is important to note that this convention is not universally accepted, and often researchers do not distinguish between sex and gender because of different philosophical viewpoints.

The importance of studying gender as a social category is to provide meaningful explanations of historical and cultural relationships between women and men and to better understand social organization as it relates to gender, including gender expectations, educational and professional opportunities, and power relations. In U.S. society, there are still large discrepancies between men and women in social roles, earning power, and occupational status. In families where both parents work full-time, the majority of household chores and child care is the responsibility of the woman. The income of the average working woman is significantly less than that of the average working man. From 1951 to 1999, 90% of university professors in chemistry, physics, mathematics, and engineering were men. In the United States, most positions of political power are held by men.

What causes these differences? There are many possible social explanations, including the notion that, in general, parents, teachers, and society expect and encourage different things from girls and boys, men and women. Gender stereotyping is pervasive: Males are more widely believed to be dominant, independent, aggressive, and achievement-oriented, whereas females are generally believed to be nurturing, affiliative, less esteemed, and more helpful in times of distress. Historically women have been oppressed and have not been offered equal opportunities. Before 1928, women could not vote. Women were less likely to be admitted to institutions of higher learning even when their qualifications were equal to those of male applicants. In fact, until 1969 many private colleges and universities did not admit women. In the past 50 years there have been some changes in social roles and economic and political status of women due to legislation and efforts to abolish various inequities; however, in most settings gender stereotypes persist and provide justification for the existing social arrangement of males as dominant and women as subordinate.

Measuring Gender Differences

To study differences between men and women, researchers must be able to measure them, but measuring psychological characteristics is difficult because they cannot be seen directly and are inferred from observable behaviors. There is no general agreement on the most appropriate instruments or methods for measuring psychological or behavioral differences. Research on gender differences is exceedingly complex because individuals have their own beliefs about gender and these widely held beliefs, or stereotypes, provide a lens through which researchers view behaviors and research findings.

In *The Psychology of Sex Differences* Eleanor Maccoby and Carol Jacklin outlined four major problems associated with studying sex differences:

1. Overreporting of significant differences
2. Influences of stereotypes on perceptions of researchers and participants
3. Sex differences being situation dependent
4. Disagreement in results using different methods

First, there is an overreporting of significant sex differences because there is a tendency to publish studies in which a sex difference is reported, whereas studies not showing significant sex differences go unpublished. Second, people tend to see what they expect to see based on personal beliefs, assumptions, and experiences. For example, a father may report his son plays rough because that is what he expects boys to do. The next problem with studying sex differences is the possibility that a sex difference is dependent on a situation as is represented in the research on achievement motivation. In certain situations girls show more achievement motivation than boys, but in other situations the sex difference disappears or reverses. The final problem is in situations where different methodologies produce conflicting results. An example is the study of anxiety and fearfulness. On self-report measures girls indicate more anxiety and fearfulness than boys, but based on teacher ratings of behavioral observations, there is no sex difference.

Although there are problems confounding the study of differences between males and females, Maccoby and Jacklin site evidence supporting significant gender differences in aggression, juvenile play behavior, and specific cognitive abilities, including visuospatial ability, mathematical ability, and verbal ability. Later findings suggest gender differences exist only in

specific subcategories of the specific abilities. The largest gender differences appear in mental rotation tasks as early as 4 years of age. Additionally, there is evidence to support gender differences in personality traits such as nurturing tendencies, dominance assertion, and activity level in children.

Gender Differences in Cognition

The gender differences in general intelligence and specific abilities have been widely studied, producing a large amount of information about female and male cognitive abilities. Most general intelligence measures show insignificant differences between the sexes. However, some subtests of the Wechsler intelligence scales show moderate differences favoring females on the symbol/coding subtest and favoring males on the information and block design subtests. A well-known cognitive difference is on the measures of visuospatial abilities. The largest difference favoring males is on measures of mental rotations. The differences range from small on two-dimensional tasks to large on three-dimensional tasks. The difference in mental rotations is apparent in childhood but seems to increase with age. Spatial perception tasks also show gender differences, with the differences appearing larger in adults than in young people.

Mathematical and verbal ability has also been studied for gender differences. Similar to visuospatial abilities, mathematical abilities vary with age and the type of ability measured. Differences are also dependent on the population studied. Overall difference in mathematical ability is negligible but in the direction of favoring females. However, in older, selected samples, such as college students taking standardized tests such as the Scholastic Aptitude Test (SAT) and the Graduate Record Exam (GRE), differences favor males on problem-solving tasks. Analyses of verbal abilities indicate a slight overall female advantage. When looking at individual tasks, males show a negligible advantage for analogies, and females show a small advantage in measures of speech. Some specific measures of verbal fluency may show larger female advantage, but other verbal abilities show essentially no sex differences.

Gender Differences in Aggression

Findings suggest greater aggression in males than females. These differences are seen in several contexts and across cultures. Males are found to be more aggressive in fantasy, administer more verbal insults, show a greater modeling of aggressive behavior, and self-report aggression to a greater extent than females. Girls are more likely than boys to engage in relational aggression including behaviors such as trying to make others dislike an individual by spreading malicious rumors or ignoring another child when angry.

The reasons behind these differences are unclear. Many scholars suggest a focus on *how* gender works to explain *why* gender works. Next, several theoretical models of gender are presented that stress how gender-related behaviors emerge or are acquired. Three categories of models are described: biological influences on gender differences, early acquisition of gender-related behaviors, and sociological influences on behavioral differences between men and women.

Biological Influences

Biological models of how gender-related behaviors emerge or are acquired argue for genetic, hormonal, and physical factors as determinates of sex differences. These models consider how male/female biological differences influence gender identity and gender roles.

Chromosomes and Hormones

Biological foundations may influence the social and psychological dimensions of being male or female. Humans have 46 chromosome pairs. The chromosome difference of males and females occurs in the 23rd pair. Females have two X chromosomes and males have an X and Y chromosome. In addition to having different chromosomes, males and females experience differing levels of sex hormones. There are two main classes of sex hormones: estrogens, influencing the development of female physical sex characteristics, and androgens, promoting the development of male physical sex characteristics. After the first few weeks of gestation, the Y chromosome in the male embryo triggers the development of testes secreting large amounts of androgens and leading to the development of male sex organs. In females, low levels of androgens allow development of female sex organs. It has been suggested that early androgens might influence later behavior, as seen in increased activity levels among boys.

Recent research suggests gender identity begins to develop at this early stage of embryonic development. Males (an X and Y chromosome) with a rare birth

defect causing ambiguous genitals who are surgically assigned to be female frequently report feeling like boys and may change their gender back to male even without knowledge of their birth history. This issue is complex because although the majority of birth males assigned to be females identified themselves as male, nearly one third of study individuals were well adjusted living as females. The nature versus nurture debate is reified with the question of how much gender identity can be attributed to genetics versus socialization.

Evolutionary Psychology

Evolutionary psychologists argue adaptation during the evolution of humans produced psychological differences between males and females. Because of differing roles in reproduction, males and females faced different pressures to survive. Natural selection favored males with short-term mating strategies because having multiple sex partners improved the likelihood males would pass on their genes. Females who devoted effort to their offspring and chose mates who provided resources and protection were favored by natural selection. Males competed to acquire more resources to access females; therefore, according to evolutionary psychologists, males evolved to be more violent, competitive, and risk taking. These evolutionary differences between males and females explain gender differences in sexual attitudes and behaviors, according to evolutionary psychologists; however, in other areas of psychology, males and females are similar.

Evolutionary psychology is based in speculations about prehistory, and critics claim that there is no evidence for these speculations. Critics additionally point out that people have choices about their behaviors, and thus, prehistoric humans could have changed behaviors that were adaptive to a particular environment. Furthermore, evolutionary views pay little attention to cultural and individual variations in gender differences.

Social Influences in Early Learning

Many social scientists theorize gender-related behaviors and psychological gender differences emerge or are acquired due to social experiences. Within these social theories, perspectives such as psychoanalytic theory, social cognitive theory, cognitive development theory, and gender schema theory emphasize the early acquisition of gender-related behaviors and assume that early learning ultimately accounts for adult gender differences in a wide variety of behaviors.

Psychoanalytic Theory

The psychoanalytic approach seeks to explore a child's gender development from Sigmund Freud's view of the nature of human sexuality. Freud believed that the preschool-age child (3–5 years of age) develops a sexual attraction to the opposite-sex parent. At 5 to 6 years of age, the child experiences anxiety due to this sexual attraction and renounces it and then identifies with the same-sex parent. Unconsciously, when identifying with the same-sex parent, the child adopts that parent's characteristics, therefore developing his or her gender behavior. However, critics argue gender development is not as Freud proposed. More recent studies show children become gender-typed much earlier than age 5 or 6, and they develop masculine or feminine traits even when the same-sex parent is not present.

Social Cognitive Theory

An alternative explanation of how children develop gender-typed behavior is the social cognitive approach. American psychologist Albert Bandura emphasized the links between cognitive processes and environment and behavior. His research focused on imitation or modeling in which learning occurs through observing what others do. In this theory, a child develops gender-typed behavior through the rewards and punishments children experience for gender-appropriate and gender-inappropriate behavior. There are socialization pressures to behave consistent with stereotypical gender norms, and these norms are gradually internalized and displayed. Children learn about gender norms from their parents, other adults in their neighborhood and in the media, and from their peers.

Parental Influences. Gender development in children is influenced by parents' actions and examples. Parents often use rewards and punishments to teach daughters to be feminine and sons to be masculine. Mothers are usually charged with the role of nurturer; however, fathers appear to play an important part in gender-role development. Fathers are more likely than mothers to act differently toward sons and daughters, contributing more to gender-role differences. However, there is

variation in the amount of gender-role pressure on individual children. Some parents encourage girls to be more nurturing and emotional than boys, for example, by reinforcing gender stereotypes in play by giving girls dolls to play with while engaging in aggressive play with their sons. Other parents make few distinctions between boys and girls. These parenting differences, in conjunction with variation in individual children's temperaments, produce a range of gender-role behaviors within children of the same sex.

Peer Influences. Although parents provide the earliest modeling of gender roles, peers also respond to and model masculine and feminine behavior, often rejecting children who display actions characteristic of the other gender. There is evidence boys experience greater pressure to conform to gender-role stereotypes than girls.

There is increasing attention to the importance of gender in peer relations. Maccoby studied the role of gender in children's peer groups. Beginning in the preschool years, children show a preference for spending time with same-sex playmates. This preference continues from 4 to 12 years of age, and during elementary school, children spend a majority of their free time with children of their own sex. Children show own-sex favoritism, but Maccoby found that only a few behaviors are differentiated by sex: Boys and girls showed differences in their playmate preferences, rough-and-tumble play, direct aggression, and themes enacted in pretend play.

Cognitive Development Theory

Cognitive development theory focuses on the personal self rather than the social self and emphasizes the active construction of individual understandings of gender. Developmental psychologists believe the goal of social development is to cultivate a self-concept distinct from others yet connected to others. From infancy to adolescence, individuals form an identity based on their psychological self and their social self. The psychological self is based on an individual's knowledge and belief in their unique characteristics, abilities, and preferences. The social self is how the individual identifies with parents, peers, and social groups. Developmental psychology maintains gender categories are meaningful for young children.

Applying the conservation and categorization skills identified by Piaget, cognitive development theory describes three levels of children's developing

understanding: awareness, identification, and constancy. First, children must develop an awareness of gender categories. Most children can correctly label gender categories shortly after their second birthday, and nearly all children can sort photos based on gender by the time they are 3 years old. In the identification stage, children can label their own gender. Findings suggest children can identify their own gender between the ages of 2 and 3 years. Lawrence Kohlberg proposed gender-typed behavior only occurs after children develop constancy, the learning that a social category is not changeable. Kohlberg suggests constancy develops by the identification of group membership, stability of group membership over time, and consistency of group membership even if there are changes in an individual's appearance or context. Therefore, if children understand the unchanging nature of being a boy or a girl, they develop gender identity and seek information about appropriate behaviors for their sex and behave according to their findings.

However, recent evidence has shown children do not develop gender constancy until they are approximately 6 or 7 years old. Gender preferences for toys, clothes, and games are often evidenced before this time. So, contrary to Kohlberg's assertions, critics argue gender-typed behavior does not appear to depend on gender constancy.

Gender Schema Theory

Similar to cognitive development theory, gender schema theory asserts individuals construct their own ideas of gender; however, gender schema theory does not include gender constancy as necessary before gender-typing of behaviors and preferences begins. Children gradually develop *schemas*, cognitive structures that are networks of associations guiding an individual's perceptions. A *gender schema* is how a child organizes the world in terms of female and male and defines what is gender-appropriate and gender-inappropriate in the child's culture. Children are internally motivated to conform to these gender schemas, and these schemas often promote stereotypical gender norms.

Sociological Influences

Sociological models propose aspects of the social structure, such as women and men filling different social roles, to explain patterns of behavioral differences between men and women. Social role theory,

social identity theory, and expectation states theory conceptualize social structure as a primary determinant of gender-linked behavior.

Social Role Theory

In most cultures, women have less power and status than men. Women control fewer resources, perform more domestic work, receive lower pay, and are not equally represented in high-level administrative roles. Alice Eagley's social role theory states that gender differences result from these contrasting roles of women and men. According to Eagley, the social hierarchy and division of labor cause gender differences in power, assertiveness, and nurturing behaviors. Because of the status differences in society, women had to adapt to roles with less power and, consequently, developed more cooperative, less dominant profiles than men.

Social Identity Theory

Social identity theory examines how people identify with a social group and how social identity leads to group phenomena such as conformity, cohesiveness, and intergroup discrimination. Social identity theory is based on the three interrelated concepts of social identity, social categorization, and social comparison. *Social identity* is a self-concept developed from belonging to a social group or groups and the value an individual places on a group membership. Once a person internalizes a group identity and incorporates the identity into his or her self-concept, group membership has personal consequences for the individual. *Social categorization* is the cognitive separation of the social environment into social categories. The categorization creates a system of the social world. Individuals can be oriented into society by defining their individual category memberships. The categories of the social world have status and often compete for resources, prestige, and power. Members of powerful groups normally have a positive social identity, and subordinate group members have a more negative social identity. Individuals seek to differentiate their own group positively from others to enhance their own self-esteem. *Social comparison* is the evaluation of groups by comparing them on related dimensions. Depending on one's evaluation of the groups to which one belongs, group membership can have a positive or negative effect on one's social identity.

Social identity theory suggests that gender as a social category will exaggerate male and female differences as individuals seek to differentiate from other groups. Furthermore, when an individual is perceived as a male or a female, the stereotypes of the group, including status, prestige, emotional experiences, goals, norms, and traits, are often accepted by the group members and influence self-perception. With males perceived as the powerful group in most cultures, based on social identity theory, males will have a more positive social identity and self-esteem than females. Also, males and females may accept more stereotypical gender norms and exhibit gender-typed behaviors when defined according to gender group membership.

Expectation States Theory

Expectation states theory focuses on small, task-oriented groups. In these groups, members evaluate their own and each other's potential contributions to the group task to determine how to reach the group's goal most effectively. Because they share cultural beliefs, group members share performance expectations for each other. Expectations for performance are based on status characteristics, attributes culturally associated with different levels of esteem and value. These characteristics can be specific skills relative to the task or diffuse characteristics such as gender, race, education, or age. When a group contains both men and women, participants tend to expect men to contribute more to the group's goal even if gender is not relevant to the task. The expectations become self-fulfilling prophecies. Because of higher expectations of them, men tend to contribute more to group discussion, behave assertively, and emerge as leaders in the group. Even when male and female group performance is identical, men's performance is evaluated more favorably. This theory provides an explanation for how societal beliefs are reproduced in group interactions and helps to explain the influence of beliefs about gender on social life.

Integrative Models

Biological influences provide explanations for gender differences focused on the action of prenatal androgens on the developing brain and differences in activity levels or play styles of boys and girls. Social theories explain the way in which gender roles can be learned either early in life or through societal

pressures. Other researchers suggest biological and social influences interact to produce complex understandings of gender. A biosocial or interactionist model suggests both biological and environmental influences are important and are likely to interact in various ways during development. There are many models of biosocial shaping of gender, and the models are often modified and replaced in light of further findings. The explanations of gender differences are not static and will continue to change as our understanding of human biology, social influences, and societal structures develops.

Promoting Gender Equity in Education

Title IX of the Education Amendments of 1972 was modeled on the Civil Rights Act of 1964 prohibiting discrimination on the basis of race, color, or national origin. Title IX states: "No person in the United States shall, on the basis of sex, be excluded from participation in, or denied the benefits of, or be subjected to discrimination under any educational program or activity receiving federal assistance." The law requires any institution receiving federal funds to maintain policies and practices that do not discriminate based on sex. To help promote gender-equitable schools, the law requires males and females to receive equal treatment. Now, more than 35 years after the passage of Title IX, there are more women in America's colleges and universities, and women receive the majority of master's degrees. Although women have made great strides, some gender differences in America's schools still exist.

Based on achievement test scores, the gap between boys and girls has closed in the United States, and females tend to earn higher grades and are the majority in undergraduate education. Attention is shifting to the underachievement of boys in U.S. schools. Boys are documented as having more discipline problems, lower grades, and higher dropout rates than girls. Boys are overidentified in special education programs and are meeting with less success in the average classroom. Underachievement of boys is an issue; however, there are still issues facing girls. For example, young women lag behind males in mathematics and science achievement in high school and in pursuing degrees in math and science in college. In addition, although women are earning college degrees in record numbers, the average income of a working woman is significantly less than a working man.

Gender equity should benefit boys and girls; it is a complex issue and it cannot be thought of as helping only boys or only girls.

Understanding what gender stereotyping is, why it is so pervasive, and how gender beliefs are developed in individuals, groups, and society can help address gender bias in the classroom. Gender is a salient social category at a young age. Children show extensive knowledge of traditional gender stereotypes as early as the toddler years, and their knowledge increases as they move into preschool. Children link specific toys, objects, activities, occupations, and clothing with males and females based on gender stereotypes, with an exaggeration of between-sex differences and within-sex similarities. Older children make complex inferences using gender-typed characteristics to predict competency, behavior, personality traits, roles, and occupations. Gender stereotypes in children are rigid and highly resistant to change. Studies show children make negative judgments about people who violate gender stereotypes, demonstrating their difficulty in accepting change to their beliefs about gender roles.

For boys and girls to have available a full range of opportunities and choices, students must not be limited by academic, athletic, or career stereotypes. Because children enter school with preexisting beliefs about gender, educators and administrators have the task of providing role models, quality educational materials, and highly qualified teachers who are sensitive to gender-based beliefs and behaviors and who can help children reflect on their gender stereotypes and reshape their beliefs about gender.

Penny L. Burge and Kimberly Filer

See also Aggression; Gender Bias; Gender Differences; Gender Identity; Schemas; Social Development

Further Readings

Archer, J., & Lloyd, B. (2002). *Sex and gender* (2nd ed.). Cambridge, UK: Cambridge University Press.

Bennett, M., & Sani, F. (Eds.). (2004). *The development of the social self.* Hove, NY: Psychology Press.

Bussey, K., & Bandura, A. (1999). Social cognitive theory of gender development and differentiation. *Psychological Review*, *106*, 676–713.

Eagly, A. H. (1987). *Sex differences in social behavior: A social role interpretation.* Hillsdale, NJ: Lawrence Erlbaum.

Hawkesworth, M. (1997). Confounding gender. *Signs*, *22*(3), 649–685.

Hines, M. (2004). *Brain gender.* New York: Oxford University Press.

Kohlberg, L. (1966). A cognitive-developmental analysis of children's sex-role concepts and attitudes. In E. E. Maccoby (Ed.), *The development of sex differences.* Stanford, CA: Stanford University Press.

Maccoby, E. E. (1988). Gender as a social category. *Developmental Psychology*, *24*, 755–765.

Maccoby, E. E., & Jacklin, C. N. (1974). *The psychology of sex differences.* Palo Alto, CA: Stanford University Press.

GENDER BIAS

Gender bias refers to the differential treatment of individuals based on their gender. Such treatment can be negative or positive and is often subtle and executed unwittingly. Although the phrase frequently describes unfair or unequal treatment of women, it can describe differential treatment of men as well. Gender bias is found in many environments and cultures. Men and women are often assigned different obligations within the family structure and receive unequal treatment in the workplace. Gender bias within educational institutions is of particular importance because of the impact such differential treatment has on child development and later life opportunities. Although great strides have been made to reduce gender bias in the classroom, gender stereotypes are still present. Gender bias in socialization, expectations, teacher interactions, and classroom resources prevent children from obtaining an equal education.

Title IX

Gender bias in education gained attention in the 1960s. Before this, classes were often segregated according to sex, and educational expectations were gender specific. Girls were expected to take home economics, while boys were encouraged to take shop, math, and science courses and participate in sports. In 1972, Congress enacted Title IX, an education amendment that prohibited sex-based discrimination in schools under penalty of loss of federal funds. Since the enactment of Title IX, education in the United States has made significant advancements toward gender equality. However, while blatant forms of gender bias have been removed from the classroom, subtle gender biases continue to exist.

Socialization and Academic Performance

At first glance, it may seem that boys are currently the endangered gender in the U.S. education system. Male students comprise 80% of high school dropouts and 66% of the learning-disabled population within U.S. schools. Compared with female students, male students struggle with reading and writing, receive a higher percentage of failing grades, and are less likely to attend college. However, substantial evidence suggests that despite the fact that females have surpassed males in many ways, girls are still discriminated against within education.

In elementary school, girls match or exceed male performance across subjects and assessments. However, by twelfth grade, most girls have fallen behind. This trend is visible in overall performance and is most pronounced in the areas of math, science, and technology. High school girls tend to perceive these courses as difficult, masculine subjects and often avoid taking them. Additionally, counselors and teachers do not actively encourage girls to pursue these areas. Without knowledge of these subjects, girls limit their career options and are effectively excluded from many lucrative career opportunities. Girls also perform less well than boys on standardized tests, such as the SAT, and consequentially receive less academic scholarships than boys.

As these gender discrepancies are not apparent in early education, they likely are due to gender differences in socialization. Modern Western standards of femininity have a negative impact on girls' self-esteem and career aspirations. Girls are evaluated on their appearance and are praised for being quiet, composed, and tidy. Beauty, popularity, and conformity become priorities for young women, often at the expense of their academic performance. In contrast, boys are taught to value independence and competence and are allowed to show more assertive and disruptive behaviors.

The impact of these subtle gender stereotype reinforcements is significant. In elementary school, girls report high self-esteem, but years of gender-biased socialization slowly erodes their confidence, self-esteem, and independence—as well as their academic performance. Moreover, socialization has led girls to

see academic success as the result of luck and academic failures as the consequences of their own deficiencies. Boys, on the other hand, tend to claim responsibility for their successes and view their failures as due to lack of effort.

Teacher Attention and Classroom Resources

These gender biases are present within student-teacher interactions as well. Overall, teachers spend less time with their female students and challenge them less often. Boys, on the other hand, are called on more often, especially for more intricate or abstract questions. Additionally, boys are praised more for their correct answers and encouraged to expand on responses. In contrast, girls' responses are more likely to be simply acknowledged. Although these differences are slight and unconscious, it leads to girls refraining from asking or answering questions or challenging the teachers' statements.

Gender bias in education can also be found within classroom resources. Many textbooks used today continue to demonstrate gender bias. Women are often absent in historical accounts, and their accomplishments are given little attention or omitted all together. When women are featured, they are often stereotyped as being submissive and obedient or put in stereotypical female roles, such as nurse or kindergarten teacher. A linguistic bias also remains in favor of men with male terminology, such as *mankind*, *policeman*, and *forefather*, prevailing throughout textbooks, whereas terminology associated with women maintain their passivity (e.g., women were *allowed* to work and were *given* equal rights).

Solutions

Because of the serious consequences that gender bias has on the development of young people, it is important that educators and institutions take steps to eliminate bias from all educational settings. The approach should be multifaceted. Educators must be aware of gender bias behaviors and be trained in strategies for altering these behaviors. Educational materials should also reflect gender equality. Teacher training should include strategies that equalize attention given to male and female students and promote gender equality in the classroom. The curriculum of all grade levels should incorporate strong male and female role models.

Most importantly, students themselves should be encouraged to see beyond society's gender stereotypes.

Cathleen Clerkin

See also Athletics; Discrimination; Gender; Gender Differences; Gender Identity

Further Readings

Chapman, A. (2002). Gender bias in education. *Multicultural Pavilion.* Retrieved October 21, 2006, from http://www.edchange.org/multicultural/papers/genderbias.html

Goetz, J. (1996, April 25). Education expert: Classroom gender bias persists. *Cornell Chronicle.* Retrieved October 15, 2006, from http://www.news.cornell.edu/Chronicle/96/4.25.96/gender.html

Owens, S. L., Smothers, B. C., & Love, F. E. (2003). Are girls victims of gender bias in our nation's schools? *Journal of Instructional Psychology*, *30*, 131–136.

Zittleman, K., & Sadker, D. (2003). Teacher education and gender equity: The unfinished revolution. *Educational Leadership*, *60*, 59–62.

GENDER DIFFERENCES

The words *gender* and *sex* are two different words with two different meanings that are, nonetheless, often confused and used interchangeably; the word *gender* is thought to be a softer, more polite way to talk about sex, particularly sexual differences between men and women. *Sex*, however, refers to the biological distinction of male and female (i.e., his sex is male; her sex is female), whereas *gender* refers to cultural aspects of masculine and feminine, the degree to which and how one understands what it means to be male or female. To illustrate, many languages (e.g., French, German, Spanish) assign gender to nouns (e.g., in Spanish, the noun *la casa* "house" is feminine, whereas *el reo* "river" is masculine). Many people believe that there are differences between men and women (e.g., men are objective, solution oriented, and dominant; women are passive, nurturing, and intuitive) and differences in the ways in which each sex relates to, and understands itself to be in, the world. Having an understanding of these differences is important for anyone working in educational psychology.

However, it is important to know that many of these differences that are assumed to be normative, universal, and timeless may, similar to language, actually be

culturally constructed and indeed vary from culture to culture and generation to generation. With only a few exceptions (e.g., visual-spatial abilities, aggressive behavior), there is little empirical evidence for many gender differences between men and women.

This entry reviews gender differences between men and women, many of which are stereotypical assumptions and a few of which have empirical support. A brief review of feminist contributions to the subject is also considered, as well as the *binary construct of gender* and some of its criticisms, including that not all cultures share Western notions of masculine and feminine behavior. Not everyone's chromosomal makeup is strictly male (i.e., XY chromosomes) or female (i.e., XX chromosomes); rather, some people are born intersex. Therefore, implications for sexual orientation are reviewed, as well as transgender, transsexual, and intersex concerns, including an illustration from the well-known John/Joan case. Multicultural considerations are also discussed, as well as applicable diagnostic considerations and their importance to educational psychology.

Feminist Critiques of Gender

Many contemporary thoughts about gender differences, especially universalized and normalized assumptions regarding gender roles, are historically rooted in feminist critiques of gender that emerged from various political and philosophical movements of the 19th and 20th centuries, including (a) the U.S. abolitionist movement, where women argued for the eradication of all forms of inequality; (b) the U.S. suffrage movement, where women disputed the notion that biological differences should be used to limit one's citizenship; (c) the influence of 20th-century French intellectual thinkers, such as Michel Foucault; as well as (d) the U.S. countercultural revolution of the 1960s and continued attempts by women throughout that era to pass the Equal Rights Amendment. These various movements built on one another, setting the historical backdrop upon which contemporary notions of gender have evolved.

One current trend in some feminist thought is to acknowledge that there are differences between men and women and the ways in which they relate to each other and the world but also to note that for most of human history, humanity has predominantly heard from only half the human species (i.e., the male half) and that women, the other half, have unique perspectives to share and contributions to make. By acknowledging and understanding these differences, examples of which are given in the following section, it is thought that men and women can better understand and relate to one another, while improving their communication with each other. Indeed, many of these differences are noted in the psychological literature having to do with relationship issues, and no doubt, it can be helpful for couples to better understand the gendered ways in which they may be relating to each other and the world.

Gender Differences and the Binary Construct of Gender

Men and women are said to be different, not only in their biological sex but also in the ways in which they understand and relate to the world and each other. Eleanor Maccoby and Carol Jacklin were among the first to study gender differences between men and women, finding that, despite a number of stereotypical assumptions (e.g., men are competitive and women are passive), there were actually only four areas in which empirical evidence supported a view of gender differences:

1. Verbal abilities

2. Visual-spatial abilities

3. Mathematical abilities

4. Aggressive behavior

These four areas lend support to the *gender similarities hypothesis* that men and women, as well as boys and girls, may be more alike than different, and within-group differences may be greater than between-group differences.

More recently, Janet Hyde completed a meta-analysis of gender differences, finding that men and women are more similar than different, with only a few exceptions. One exception is with motor performance, especially with throwing distance and velocity after puberty when men tend to have larger muscles and bone mass than women and can therefore throw objects farther and faster than women can. Another interesting area of gender difference has to do with sexuality: Men tend to have more incidences of masturbation than do women, as well as more casual attitudes regarding uncommitted sexual relationships. Finally, men are known to be more aggressive than

women, especially regarding physical violence, and there is an abundance of crime statistics to support this view. However, it is also known that women can be equally aggressive, especially in relationships, although it should be noted that women are more likely to be sexually victimized, seriously injured, or murdered by their partners than are men.

Still, a number of authors have supported the gender differences hypothesis, which tends to be very popular, not only within the media but also among the general public. John Gray's bestseller *Men Are From Mars, Women Are From Venus*, for example, notes that men are often thought to be thinkers, leading with their brains, whereas women are thought to be feelers, leading with their hearts and intuitions. Men are said to be *fix it* oriented and solution focused, whereas women offer advice and give direction. Men often do not want to talk about their feelings and will retreat in silence when something is bothering them, whereas women will want to talk and process things that are on their minds, if not with their partners, then with other women. Men like to feel needed, useful, and productive, while women want to feel loved, cherished, and respected. An insightful understanding of these general differences can be helpful, and they point toward what could be called an *essentialist view* of gender differences that are assumed to be universal and normative differences between men and women.

Social construction theory, however, holds that these generalizations of gender differences may be reflective of Western perspectives that are socially constructed, learned, and then universalized and normalized to others. Furthermore, they promote a duality, a binary understanding of male and female, masculine and feminine, in which male equals masculine, and thus, to be a man, a *real* man, is to be masculine (e.g., interested in sports, while fixing problems and avoiding feelings), whereas to be a woman, a *natural* woman, as Aretha Franklin sings (and Judith Butler notes), is to be feminine, and to be feminine is to be that which is *not* masculine. Herein lays a key point regarding binary constructions: They are often oppositional and hierarchical in which each dyad defines itself in opposition to the other and generally one dyad is dominant over the other.

Thus, to be masculine (e.g., dominant) is to *not* be feminine (e.g., submissive). Or to use another illustration, to be heterosexual (e.g., dominant) is to *not* be homosexual (e.g., submissive). Binary constructions are generally thought to be culturally derived and maintained (e.g., little boys are often dressed in blue, a culturally masculine color, and encouraged with masculine rough-and-tumble play, *not* Barbie dolls, the latter for little girls dressed in feminine pink). Binary constructions create them versus us, either/or dualities in which clients and students may find themselves trapped unknowingly. For example, a young man interested in pursuing gymnastics or dance, both often thought to be more feminine than masculine pursuits, may hesitate to develop his interests and talents, for fear of being perceived as feminine, that is, not masculine, and if not masculine, then perhaps not heterosexual, regardless of his actual sexual orientation or any other typically masculine interests he may have.

Similarly, a young woman may feel trapped in culturally derived gender roles and expectations. Psychologists and others working with clients and students can help both men and women deconstruct and evaluate the societal scripts and gender roles they may feel pressured to pursue, while also finding examples of role models who broke free of culturally derived gender expectations. For example, a woman interested in becoming a pilot, a traditionally masculine occupation, might find inspiration in knowing that Amelia Earhart was the first pilot to break many records for both aviation and women.

Problems With the Binary Construct of Gender

There are several issues that render problematic the binary construction of gender, including the way it often assumes, universalizes, and normalizes Western understandings of masculine and feminine, when in fact, not all cultures adhere to Western notions of masculine and feminine behavior. For example, the well-known anthropologist Margaret Mead studied indigenous tribes in New Guinea, finding that gentleness was the ideal temperament for both men and women, whereas in another tribe, aggressiveness was more highly appreciated for both men and women, and yet in a third tribe, dependence and affection were ideal for men, while aggression and dominance were valued in women. In another study, women in 57 different societies were found to be responsible for carrying heavy loads, although men are generally thought to carry heavy loads. Women are often thought to be responsible for cooking and sewing, and yet in some societies, men do the cooking and sewing.

Building and construction is often thought to be masculine work, and yet, in at least 14 different cultures, women do the building. Such differences among cultures call into question universalized and normalized assumptions regarding gender differences between men and women.

Additionally, the binary construction of gender is problematic because it assumes that all people are either male or female, when in fact, not everyone fits into either an XY (i.e., male) or XX (i.e., female) chromosomal duality. Some estimates note that one out of every 2,000 births are believed to be intersex (e.g., hermaphroditism, ambiguous genitalia, additional X or Y chromosomes, inner testes, ovo-testis, as well as conditions such as congenital adrenal hyperplasia, androgen insensitivity syndrome, Turner syndrome, and Klinefelter syndrome). Here, the focus is only on biological sex, specifically, and yet the discussion can quickly jump to gender, and from gender, to sexuality, as parents and physicians are often pressured to make decisions and surgical alterations that assign sex to a newborn infant when the child's gender may not develop until later in life. These surgeries can result in psychological distress as the child matures, as well as physical loss of erotic sensation after the child matures. It should be noted that a growing number of physicians, in consultation with psychologists and psychiatrists, are now encouraging parents to delay surgically assigning sex to an intersex infant until the child grows older, can better understand his or her gender, and thus choose the sex that best fits that gender.

The John/Joan Case

One of the best-known cases regarding gender and surgical sex-assignment involved a heterosexual boy who, in 1967, lost his penis at 8 months due to a circumcision accident. The boy's penis was burned beyond repair when an electrocautery needle used to control bleeding malfunctioned. Desperate, the boy's parents consented to the best medical advice they could find at the time and allowed the infant to undergo a sex-reassignment surgery at 22 months of age, followed by hormonal treatments. The child was raised as a girl, while his identical twin brother was raised as a boy, and thus, the twin brother served conveniently as an experimental control.

Throughout childhood, however, it became increasingly evident that something was wrong. Not only did the child feel like a boy and have tomboyish mannerisms, but her classmates and even her teachers knew something was different. After repeated taunts in school and various suicide attempts, the child's parents told her, at age 14, what happened. She immediately demanded another sex-reassignment surgery, eventually married a woman, and adopted her children. Sadly, the story ended in 2004 when he committed suicide at the age of 38. The case is complex and sheds light on many aspects of gender identity, including the belief that, while gender norms may be culturally constructed, sexual identity may have a neurological component, and thus a person's brain may tell him or her that he or she is one sex, while the person's genitals may suggest another.

Transgender and Transsexual Considerations

Transgender refers to people who understand their gender identity to be different than their biological sex. Transgender people often describe themselves as being "trapped in the wrong body," and it should be noted that they can be gay, lesbian, bisexual, or even heterosexual. For example, a biological male who understands himself to be female may be attracted to other men and might, therefore, understand herself to be heterosexual. It is appropriate to use pronouns for transgender people that describe their gender identity rather than their biological sex.

Although it is difficult to obtain accurate statistical information, suicide rates among transgender people are thought to be especially high, even as high as 50%, because transgender people are often rejected by their families, as well as by society. Violent assaults and brutal murders have also been committed against transgender people, and sadly, when transgender people are murdered, they often are shot or stabbed repeatedly, riddled with bullets, or bludgeoned beyond recognition. Crimes against transgender people reveal society's misunderstanding and disdain for these individuals.

Thus, psychologists have an educational imperative to help overcome cultural ignorance and prejudice that contribute to stigma, low self-esteem, and high suicide and homicide rates within the transgender community. Still, despite these obstacles, many transgender people succeed, some electing to have sex-reassignment surgery and receive hormonal treatments, often at great financial expense, as well as what is thought by some

to be intrusive and require psychological counseling and inquiries (i.e., the presumption being that they must be *crazy* if they want to change their sex).

Generally, it is common to refer to transgender people who are transitioning from one sex to another with the acronyms FTM (female to male) or MTF (male to female), as applicable. Although transgender people have their own community, they are also considered to be part of the GLBT (gay, lesbian, bisexual, transgender) community and are sometimes also referred to as *queer*, a previously pejorative term that was reclaimed by some within the GLBT community as a word of empowerment, although the word can be regional in use and still considered offensive to some people.

Additionally, the word *transsexual* was historically used in reference to those who surgically changed their sex, whereas the word *transgender* was used in reference to those who understood their gender to be different from their biological sex but who did not undergo a sex-reassignment surgery. Today, the two words are often used to distinguish those whose sexual identity is different from their biological sex (i.e., transsexual) and those whose gender identity is different from their sexual identity (i.e., transgender), regardless of whether an individual undergoes sex-reassignment surgery. For example, a person born male may understand her sexual identity to be female, and her gender identity may also be feminine. In this case, both her sexual identity (female) and her gender identity (feminine) are in agreement, whereas it is her sexual identity (female) that is *transversing* her biological sex (male). Thus, she may identify as transsexual rather than transgender. Although some in the GLBT community use the word *transgender* as an umbrella term to include transgender, transsexual, and intersex individuals, it is important to note that not all transsexuals understand themselves to be transgender, and many intersex people may be transversing neither sex nor gender but rather living the sex and gender to which they were born and understand themselves to be.

Multicultural Considerations

There are a number of multicultural considerations for counselors to keep in mind when working with racially and ethnically diverse populations, especially the ways in which culture can influence gender identity. For example, *familismo* (family unity) is a traditional value that can be found among many Hispanic and Latina/o Americans, in which respect and loyalty for one's family is important. In terms of gender, Hispanic and Latino men are often expected to be strong and to be providers of the family (i.e., *machismo*), whereas Hispanic and Latina women are often expected to be nurturing, self-sacrificing, and subordinate to men (i.e., *marianismo*).

Additionally, racial and ethnic minorities may have dual identities in which gender can be an interacting influence. For example, similar to Hispanic/Latino machismo, African American men are also often expected to be strong and dominant, and yet a gay African American man may be perceived as weak and/or feminine, especially if he takes a receptive or submissive role when having sex with another man. Thus, many African American men who have sex with other men may not self-identify as *gay* or *homosexual*, two terms that are considered to be Caucasian by some African Americans. It is common to use the acronym MSM (men who have sex with men), rather than *gay* or *homosexual*, when referring to these men, many of whom may be married to women, self-identify as heterosexual, and may understand themselves to be especially masculine if they take an inserting or dominant role when having sex with another man.

Similarly, ethnic minority women may have triple identities to manage in which gender can be an influential variable (e.g., African American, female, and lesbian). It is not uncommon for multiple identities to compete with each other (e.g., identifying more as an African American than GLBT) or dualistically exclude each other (e.g., the notion that one cannot be African American and GLBT or that the GLBT community is primarily a Caucasian community). Thus, counselors should consider any multiple identities and cultural influences of gender that clients may be managing.

The transgender/transsexual community also has a number of unique multicultural considerations for counselors to be aware of, such as the *hijras* of India, Pakistan, and Bangladesh. Hijras are male-to-female transgender/transsexual or intersex individuals, often of the Muslim or Hindu faith. Similarly, contemporary transgender/transsexual people may find commonalities with ancient shamanic practices among the Inuit, as well as the Native American *berdache* or *two spirit people*, who were thought to possess both male and female spirits within their bodies and were often highly regarded within their tribes. It should be noted, however, that the term *berdache* is considered inappropriate by some Native Americans because of its French origins. Additionally, the *kathoey* or *ladyboys* of Thailand

are well known and generally accepted within their culture as either very effeminate gay men or male-to-female transgender/transsexual people. Having a global and historic perspective on transgender/transsexual issues can help counselors increase their multicultural knowledge and competence when working with this often underserved and misunderstood population.

Diagnostic Considerations

Several disorders in the *Diagnostic and Statistical Manual of Mental Disorders* (Fourth Edition, Text Revision) (*DSM–IV–TR*) may be especially relevant for clinicians to consider when working with clients who present with concerns regarding gender differences, including gender identity disorder, transvestic fetishism, as well as various eating, depressive, anxiety, and relational disorders.

Gender identity disorder is characterized by (a) a very strong and persistent desire to be, or insistence that one is, the opposite gender or sex, as well as (b) persistent discomfort about one's current sex or gender role or insistence that one's current sex or gender role is incorrect. Both criteria must be met for diagnosis, as well as evidence of significant distress or impairment in functioning. The diagnosis is not appropriate for intersex individuals. In children, boys with the disorder may exhibit a preoccupation with typically feminine activities, such as playing house, playing with Barbie dolls, or dressing in improvised women's clothes (e.g., towels, aprons, scarves), while avoiding rough-and-tumble play and sometimes sitting to urinate, perhaps pushing their penis between their legs and pretending not to have it. Girls will often prefer to wear boy's clothes and exhibit strong negative reactions to parental attempts to have them wear dresses; they often prefer rough-and-tumble play and have little use for dolls, cosmetics, or other typically feminine items.

Adults with the disorder are generally preoccupied with a strong desire to be the opposite sex and may exhibit behavior typical of that sex (e.g., attire and mannerisms). They may spend considerable time in private perfecting their appearance as the opposite sex, while also venturing out in public as the opposite sex; they may also seek hormonal treatments or sex-reassignment surgery to bring into congruence their biological sex with their sexual or gender identity. Depending on their level of development, adolescents with the disorder may exhibit characteristics of either children or adults. Psychologists should be mindful

that individuals with this disorder may be socially isolated, rejected by their families, and have low self-esteem or depression; clinicians should also be familiar with the Harry Benjamin International Standards of Care for Gender Identity Disorders.

Transvestic fetishism is found in heterosexual males and involves cross-dressing in women's clothes as a form of sexual arousal. The cross-dressing may consist of only one or two feminine undergarments underneath masculine clothes, or it may consist of an entire wardrobe of feminine attire, including wearing cosmetics. These men may be married to women, and in some marriages, the wives are aware of their husband's fetish and have accommodated it within the relationship. Transvestic fetishism, however, should not be confused with transgender or transsexual concerns, the former being a heterosexual male phenomenon regarding sexual arousal and the latter two issues having to do with gender and sexual identity, respectively, as discussed previously.

Psychologists should be mindful that low self-esteem, distress, depression, and/or anxiety, as well as relational problems may be comorbid with either gender identity disorder or transvestic fetishism. Additionally, counselors may wish to consider the ways in which media and advertising portrayals of gender differences may influence their clients' gender identity and screen for eating disorders if appropriate (e.g., anorexia nervosa, bulimia nervosa).

Psychologists and educators can better serve clients and students by being familiar with the many aspects of gender differences, including their own biases, as well as various multicultural and diagnostic considerations (e.g., *familismo* within Hispanic/Latina/o culture and gender identity disorder). Gender differences are complex, and there are many variables to consider, each influencing what it means to be male or female, masculine or feminine, or perhaps even male and feminine or female and masculine.

Andrew D. Reichert

See also Sexual Orientation

Further Readings

Butler, J. (1999). *Gender trouble: Feminism and the subversion of identity* (10th anniversary ed.). New York: Routledge.

Colapinto, J. (2000). *As nature made him: The boy who was raised as a girl*. New York: HarperCollins.

Cornell, D. (2004). Gender in America. In N. Tazi (Series Ed.) & L. Waldman (Vol. Ed.), *Keywords: Gender* (pp. 33–54). New York: Other Press.

Gilligan, C. (1982). *In a different voice: Psychological theory and women's development.* Cambridge, MA: Harvard University Press.

Gray, J. (1992). *Men are from Mars, women are from Venus: A practical guide for improving communication and getting what you want in your relationship*s. New York: HarperCollins.

Hyde, J. S. (2005). The gender similarities hypothesis. A*merican Psychologist*, *60*, 581–592.

Maccoby, E. E., & Jacklin, C. N. (1974). *The psychology of sex differences.* Stanford, CA: Stanford University Press.

Slattery, P. (2006). *Curriculum development in the postmodern era* (2nd ed.). New York: Routledge.

Spelke, E. S. (2005). Sex differences in intrinsic aptitude for mathematics and science: A critical review. *American Psychologist*, *60*, 950–958.

GENDER IDENTITY

Gender identity is commonly defined as an individual's sense of being a male or a female. For many, gender identities are aligned with physical sex characteristics including hormones, chromosomes, genitalia, and secondary sex characteristics, as well as with sex assignment and gender roles. For others, gender identities do not match one or more sex or gender traits. Several variations on this concept of gender identity exist. For example, the social identity perspective defines gender identity as the degree of awareness and acceptance of one's own gender category. Alternatively, symbolic interactionists consider gender identities as self-concepts based on fulfillment of gender roles.

Multidimensional models of gender identity illustrate the complexity and breadth of gender identity. For example, Susan Egan and David Perry offer a model of gender identity including five components: (1) knowing one's gender category membership, (2) feeling similar to others in one's category, (3) feeling satisfied with one's gender assignment, (4) feeling pressure to conform to gender roles, and (5) believing that one's gender category is superior to the other category.

Several stage models of gender identity development connecting gender knowledge to gender behavior have been proposed by researchers such as Lawrence Kohlberg, Warren O. Eaton and Donna von Bargen, and Phyllis Katz. Most models suggest that gender identity development is initiated by acquiring the ability to label the self and others with correct gender labels. Next, children learn that identity is stable throughout development and is permanent regardless of the influences of culture, motive, or hairstyle or clothing changes. Katz proposed that sexual gender, developing during puberty, and reproductive gender, developing during adulthood, also contribute to gender identity. To date, research has failed to clearly support or refute any of these models.

Evidence supports both biological and sociocultural influences on gender identity. While sex may be defined by sex chromosomes (XY for male and XX for female), sex chromosomes direct gonads to develop into testes or ovaries, which, in turn, produce the hormones between weeks 8 and 24 of gestation that influence gender development. Prenatally, testes produce testosterone, whereas ovaries do not produce considerable amounts of the hormone. Studies of girls with a congenital adrenal hyperplasia (CAH) illustrate the effects of prenatal hormones on gendered behaviors. CAH is a condition in which fetuses are subject to large quantities of hormones due to synthetic steroids. Girls (XX) with CAH are born with masculinized genitalia and undergo hormone treatment and surgical feminization, whereas boys (XY) with CAH are born with normal genitalia. As children, boys with CAH exhibit play behavior and interests similar to non-CAH boys, whereas girls with CAH exhibit play behavior and interests that are more masculine than that of non-CAH girls and slightly more feminine than CAH and non-CAH boys. Considering postnatal hormonal and environmental influences on CAH and non-CAH girls were similar, these behavior differences may be attributed to prenatal hormonal environments.

A second surge of hormones from the testes occurs shortly after birth, from the first to sixth month, increasing testosterone in boys but not in girls. These early postnatal hormones also influence gender identity development. In animal studies, when male animals are castrated and female animals receive hormones during this critical period, sex-typed behaviors are completely reversed in adulthood. Additionally, brain structure may be connected to gender identity. For example, male to female (MTF) transsexuals have been observed to have a smaller subsection of the hypothalamus than a male control group.

Biology also interacts with sociocultural factors to influence gender identity. Upon birth (and, with technology, upon ultrasound results), the presence or

absence of external genitalia signals to adults that they should dress the baby in blue, bounce the baby, and proclaim it handsome or that they should dress the baby in pink, soothe it, and praise its prettiness. Social modeling and media portrayal along with gender-related sanctions and opportunities exert influence on children's identities and gender expressions. Additionally, children play active roles in developing gender identities by displaying preferential attention to the behaviors of children of the same sex as early as 12 to 24 months for boys. Once learned, continual engagement in gendered behaviors helps maintain gender identity.

In addition to the traditional gender labels "man" and "woman," new labels have emerged as variant gender identities have gained recognition. *Transgender*, or *trans*, refers to those whose gender identity or presentation deviates from the norm. A transgender person who has changed pronouns, names, or clothing, or has undergone hormone therapy or sex-reassignment surgery to live as a member of the opposite sex, may be labeled *transsexual*. Transsexuals who were born biologically female and identify as males are *female to male transsexuals* (*FTM*, or *transmen*) while those who were born biologically male and identify as females are *MTF transsexuals* (or *transwomen*). A person who identifies as *genderqueer* transgresses gender norms or renounces the two-gender system. Gender-related labels also vary by community. Whereas queer communities may use the terms *butch* and *femme*, straight communities may use the terms *tomboy* and *girly-girl*. Furthermore, some individuals or communities may reject the use of gender-related labels altogether.

Cross-gender identification along with discomfort with one's assigned sex was officially recognized as a psychological disorder, gender identity disorder (GID), in 1980, when the American Psychiatric Association included GID in the third edition of the *Diagnostic and Statistical Manual of Mental Disorders*. Labeling transsexuals as having a psychological disorder has proven controversial. Critics, including psychologists Darryl Hill and Kelley Winters, argue that transsexual people suffer psychological distress from parents and a society that fail to accept nonstereotypical gender expression but are otherwise healthy and well-adjusted. These critics suggest replacing GID with *gender dysphoria*, a term describing persistent distress with sex characteristics or prescribed gender roles, or altogether removing diagnoses related to gender identity. In contrast, proponents, including psychiatrists

Robert Spitzer and Paul J. Fink, argue that cross-cultural expectations to fulfill gender roles associated with biological sex lend legitimacy to the identification of transsexualism as a disadvantageous psychological disorder. Future research regarding the origins, development, and mutability of gender identity may clarify this issue.

Anne S. Beauchamp

See also Androgyny; Gender; Sexual Orientation

Further Readings

Boston Women's Health Book Collective. (2005). *Our bodies, ourselves* (4th ed.). New York: Simon & Schuster.

Eagly, A. H., Beall, A. E., & Sternberg, R. J. (Eds.). (2004). *The psychology of gender* (2nd ed.). New York: Guilford Press.

Frable, D. E. S. (1997). Gender, racial, ethnic, sexual and class identities. *Annual Review of Psychology*, *48*, 139–162.

Hausman, K. (2003). Controversy continues to grow over *DSM*'s GID diagnosis. *Psychiatric News*, *38*, 25.

GENERALIZABILITY THEORY

Generalizability (G) *theory* is a statistical theory for evaluating the dependability ("reliability") of behavioral measurements. G theory pinpoints the sources of systematic and unsystematic measurement error, disentangles them, and then estimates their magnitudes simultaneously.

In G theory, a behavioral measurement (e.g., an academic self-concept score) is conceived of as a sample from a *universe of admissible observations*. This universe consists of all possible observations that decision makers consider to be acceptable substitutes (e.g., scores sampled on Occasions 2 and 3) for the observation in hand (scores on Occasion 1). Each characteristic of the measurement situation (e.g., survey form, item, occasion, rater) is a potential source of error and is called a *facet* of a measurement. The universe of admissible observations is defined by all possible combinations of the levels (called *conditions*) of the facets. To evaluate the dependability of behavioral measurements, a *generalizability* (G) *study* is designed to isolate and estimate as many facets of measurement error as are reasonably and economically feasible.

For example, consider a G study in which all persons in a sample respond to the same randomly sampled 10 academic self-concept items on the same two randomly sampled occasions. In this G study, the facets of the measurement are items and occasions. As the self-concept inventory was designed to capture systematic variation among persons, the *object of measurement*, persons, is not a source of error and, therefore, is not a facet.

An observed score for a particular person on a particular item and occasion is decomposed into an effect for the grand mean, plus effects for the person, the item, the occasion, each two-way interaction, and a residual (three-way interaction plus unsystematic error). The distribution of each component or "effect," except for the grand mean, has a mean of zero and a variance σ^2 (called the *variance component*). The variance component for the person effect is called the *universe-score variance*. The variance components for the other effects are considered error variation.

An estimate of each variance component can be obtained from an analysis of variance (or other methods, such as maximum likelihood). The relative magnitudes of the estimated variance components provide information about systematic differences in self-concept among persons (universe-score variance) and sources of error influencing the measurement. Statistical tests are not used in G theory; instead, standard errors for variance component estimates provide information about sampling variability of estimated variance components.

G theory distinguishes a *decision* (D) *study* from a G study. A D study uses variance-component estimates from a G study to design a measurement procedure that minimizes error for a particular purpose.

In planning a D study, the decision maker defines the universe that he or she wishes to generalize to, called the *universe of generalization*, which may contain some or all of the facets and their conditions in the universe of admissible observations. The D study imports G study information about selected variance to evaluate alternative designs. As decisions usually will be based on the mean (or sum) over multiple observations (e.g., questionnaire items) rather than on a single observation (a single item), D-study designs will change the numbers of conditions of the relevant facets, thereby affecting "reliability"; the more conditions there are, all else being equal, the higher the reliability will be.

G theory recognizes that the decision maker might want to make two types of decisions based on

a behavioral measurement. A *relative* ("norm-referenced") *decision* focuses on the rank order of persons; an *absolute* ("criterion-referenced" or "domain-referenced") *decision* focuses on the level of performance, regardless of rank. Error variance is defined differently for each kind of decision. To reduce error variance, the number of conditions of the facets may be increased (analogous to the Spearman-Brown prophecy formula in classical test theory and the standard error of the mean in sampling theory). G theory provides reliability-like summary coefficients for each decision: a Generalizability (G) Coefficient for relative decisions and an Index of Dependability (Phi) for absolute decisions.

Generalizability theory allows the decision maker to use different designs in G and D studies. Although G studies should use crossed designs, whenever possible, to estimate all possible variance components in the universe of admissible observations, D studies may use nested designs for convenience or to increase estimated generalizability.

G theory is essentially a random effects theory. Typically, a random facet is created by randomly sampling conditions of a facet. A fixed facet arises when the decision maker (a) purposely selects certain conditions of a facet and is not interested in generalizing beyond them, (b) finds it unreasonable to generalize beyond the levels observed, or (c) defines a small universe and includes all its conditions in the measurement design. G theory typically treats fixed facets by averaging over the conditions of the fixed facet and examining the generalizability of the average over the random facets. Alternatives include carrying out a separate G study within each condition of the fixed facet, or a multivariate analysis with the levels of the fixed facet comprising a vector of dependent variables.

Table 1 gives estimated variance components from a fully crossed, random G study of self-concept (persons × items × occasions; item scores range from 1 to 5). The large $\hat{\sigma}_p^2$ (1.256, 34% of the total variation) shows that persons in the sample differed systematically in their self-concept. The other estimated variance components constitute error variation. The nonnegligible $\hat{\sigma}_i^2$ (5% of total variation) shows that items varied somewhat in average level of self-concept reported. The large $\hat{\sigma}_{pi}^2$ (21%) reflects different relative standings of persons across items. The small $\hat{\sigma}_o^2$ (1%) indicates that the average self-concept score was stable across occasions. The nonnegligible $\hat{\sigma}_{po}^2$ (7%) shows that the relative standing of persons differed somewhat across

Table 1   Estimated Variance Components in a Generalizability Study of Academic Self-Concept ($p \times i \times o$ Design)

Source	Estimated Variance Component	Estimate	% of Total Variability
Person (p)	$\hat{\sigma}_p^2$	1.256	34
Item (i)	$\hat{\sigma}_i^2$	0.185	05
Occasion (o)	$\hat{\sigma}_o^2$	0.037	01
$p \times i$	$\hat{\sigma}_{pi}^2$	0.776	21
$p \times o$	$\hat{\sigma}_{po}^2$	0.259	07
$i \times o$	$\hat{\sigma}_{io}^2$	0.001	00
$p \times i \times o,e$	$\hat{\sigma}_{pio,e}^2$	1.182	32

occasions. The very small $\hat{\sigma}_{io}^2$ (0%) indicates that the rank ordering of item scores averaged over persons was similar across occasions. Finally, the large $\hat{\sigma}_{pio,e}^2$ (32%) reflects the varying relative standing of persons in self-concept across occasions and items and/or other sources of error not systematically incorporated into the G study.

Because more of the error variability in self-concept scores came from items than from occasions, changing the number of questionnaire items will have a larger effect on the estimated variance components and generalizability coefficients than will changing the number of occasions. For example, the estimated G and Phi coefficients for 4 items and 2 occasions are 0.73 and 0.70, respectively; the coefficients for 2 items and 4 occasions are 0.68 and 0.64, respectively. Choosing the number of conditions of each facet in a D study, as well as the design (nested vs. crossed, fixed vs. random facet), involves logistical and cost considerations as well as issues of dependability.

Noreen M. Webb and Richard J. Shavelson

See also Measurement; Reliability

Further Readings

Brennan, R. L. (2001). *Generalizability theory.* New York: Springer.

Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements.* New York: Wiley.

Shavelson, R. J., & Webb, N. M. (1991). *Generalizability theory: A primer.* Newbury Park, CA: Sage.

GIFTED AND TALENTED STUDENTS

Gifted and talented students have been the focus of research by scientists, educators, and researchers in special education, educational psychology, and general education for over a century. Although there appears to be consensus regarding the notion that there are students who are gifted, students who are talented, and students who are both gifted and talented, the same type of unity does not exist regarding the definition of gifted and talented students and how to address the needs of these students. Historically, the term *gifted* has been perceived as synonymous with high intelligence as measured by an intelligence test. The term *talented* has been viewed more subjectively, measured qualitatively, and viewed as more inclusive. As the field of gifted and talented education has developed over the years, the definitions of *gifted* have changed. The definition has evolved from a very narrow focus to a more inclusive definition and has led to a broader range of ways to identify and serve gifted and talented students. In the 1900s, Lewis Terman defined *gifted* as the top 1% level in general intellectual ability. In 1972, Sydney Marland defined *gifted and talented students* as those who are capable of high performance in general intelligence, specific academic aptitude, creativity, leadership, visual and performing arts, and psychomotor ability.

Interest in the identification of gifted and talented students is not exclusively a modern occurrence. The history of gifted and talented education can be traced to Plato, who proposed the idea of selecting children of high potential and grooming them for leadership roles. In A.D. 800, Charlemagne realized that there were poor children of promising potential who could be educated as leaders. In the 15th century, a Turkish sultan searched his empire for the most intelligent boys. Comenius, in the 17th century, wrote about individuals from poor homes who had an unusual ability for learning. During the 18th century, Thomas Jefferson advocated the radical view of providing a university education for the most intelligent, at public expense.

Before the 20th century, the concept of giftedness was equated with the genius or the child prodigy of extraordinary talents and achievement. Children who demonstrated great power of memory or possessed mature writing skills, child prodigies who received advanced college degrees at a young age, and children

who could speak several languages were displayed and asked to perform adult tasks.

The industrial era of the 19th century was not characterized by tangible national recognition of the gifted and talented. This period marked the emergence of the efforts to measure individual differences, such as the scientific studies of intelligence by Francis Galton. It was Galton's investigation into the relationships between inheritance and intelligence that led to the emergence of the theory of fixed intelligence. Building on the findings of Galton, Lewis Terman's studies solidified the theory of fixed intelligence. Terman's work dispelled the myth that the intelligence of gifted and talented students is accompanied by a wide range of deficiencies. Terman's research also supported the finding that gifted students are superior in all categories of development. This conclusion led to the belief that gifted students will survive without the help of professional adults and that they do not need special curricular and instructional support.

Contributions by Maria Montessori, J. P. Guilford, Lev Vygotsky, and Jean Piaget raised questions regarding the premise of fixed intelligence. Montessori assumed that intelligence could be developed and was not fixed. In his model, Guilford expanded the parameters of intelligence, described intelligence as educable, and concluded that there are 120 different aspects of intelligence. Vygotsky's research conclusions addressed intelligence and educational remediation, early stimulation and language development. Piaget researched the interaction between genetics and the environment as he developed the stages of cognitive development.

Continuing efforts by Jerome Bruner and Benjamin Bloom led to further support for an expanded view of intelligence. Jerome Bruner developed a theory of instruction designed to release the capacity of the mind. Bloom reexamined research data to uncover that children's early years are optimum for learning and expanding intelligence.

In the 1980s, Robert Sternberg's triarchic concept of intelligence and Howard Gardner's theory of multiple intelligences, in very different ways, emphasized the interactive quality of intelligence and the environment.

The relationship between gifted and intelligence is historical in origin and has had a major impact on how gifted students are viewed. Despite the preponderance of evidence and authorities who support both the fixed view of intelligence and the interactive concept of intelligence, researchers and practitioners in education and psychology stand on both sides of the issue. The nature versus nurture debate reigns and colors efforts to define, identify, and design programs for gifted and talented students.

Movement Toward a New View of the Gifted

Gertrude Hildreth challenged psychologists to begin viewing the gifted as individuals. She emphasized that efforts to provide lists of characteristics of the gifted may lead to stereotypes of the gifted. Gifted children are, first of all, individuals and have both strengths and weaknesses. Few are uniformly superior in all categories. Ruth Strang questioned the issue of describing the gifted in terms of a general intelligence test score. Strang suggested the use of question marks following a gifted characteristic rather than a period. Strang raised a series of questions related to whether a particular child is a prodigy or a genius; highly gifted in abstract verbal ability; academically talented, evidencing unusual ability in math; demonstrating a special skill with mechanical things; talented in art, drama, music, or dance; or socially gifted. Strang's questions reinforced Hildreth's plea to begin looking at the gifted as individuals. Recognition of the dual-language gifted, the learning-disabled gifted, the underachieving gifted, the attention deficit hyperactivity disorder (ADHD) gifted, and the culturally different gifted have all contributed to the modification of the concept of giftedness.

Characteristics of the Gifted

In delineating characteristics of gifted children, Barbara Clark has listed a series of questions that educators should ask themselves as they observe individual students. No gifted student will have all of these characteristics. Gifted students may ask many questions, attempt to do many things in innovative and unusual ways, retain a great deal of information, and be sensitive to discrimination or inequities. Gifted students may also be deficient in certain basic skills and may dislike repetitive activities. At the same time, gifted students may have great zeal for learning and may become agitated when asked to transition from their preferred area of interest. Gifted students are often talkative, perfectionistic, loners, extremely social, or very critical of others, but not always. There are certain generalizations that tend to be characteristic of

the gifted, but few traits are descriptive of all gifted and talented students.

It has been recognized that many of the positive characteristics present concomitant problems for teachers and gifted students. For example, gifted students who have a large vocabulary, facility of expression, and breadth of information may read inappropriate material for their age. They may escape into verbalism, appear to be a "know-it-all," and appear conceited to their agemates. These behaviors can cause a gifted child to be alienated by other children and by teachers. Teachers who have been trained in the appropriate instructional strategies to use with gifted children can avoid such problems by providing the gifted child with appropriately challenging reading material, teaching the child to translate his or her abstract ideas in concrete ways, and providing the child opportunities to interact with his or her intellectual peers.

The needs of highly intellectually gifted children with IQs of 145 and above may be met best by using out-of-school resources, such as concurrent enrollment and mentors. Acceleration options, independent study options, and research opportunities should be available for the gifted student. Gifted students also need opportunities to engage in real-world projects and opportunities to interact with other gifted students.

Identification of the Gifted and Talented

Well-designed identification procedures include the use of qualitative measures and quantitative measures that have been shown to be effective in uncovering the gifts of the hard-to-identify gifted and talented student. Teacher nominations also are used in identifying students for gifted programs. Teachers who have been educated to recognize giftedness are more effective in identifying students than are teachers who have not been trained. Teachers trained in the principles of gifted education are aware that gifted students can be deficient in basic skills, manifest behavior problems in the classroom, be introverts who never turn in classwork, or be students who make good grades and never cause any problems. Gifted students can also be students who excel on standardized tests or any quantifiable measure but make barely passing grades in the classroom. Academically talented students may have the best classroom grades but mediocre scores on standardized tests.

Qualitative measures include learning style inventories, informal achievement tests, questionnaires, interviews, checklists, structured observations, familial characteristics, criterion-referenced instruments, sociometric instruments, portfolios of student work, and simulations.

In an effort to identify students for gifted programs, it also should be remembered that there are varied aspects of giftedness. Quantifiable measures may not be the most effective measures to use in identifying students who are gifted in visual arts, performing arts, or leadership. Identification procedures should be as diverse as giftedness is and should always address the goals and objectives of the program or services offered. To recognize the varied dimensions of giftedness, educators should consider the impact culture can have on how giftedness may appear. The values that come with living in poverty and the cultural differences of gifted and talented children whose native tongue may be a language other than English all influence the perception and appearance of gifted characteristics. Identification decisions based on a preponderance of data drawn from varied sources characterize a strong identification model.

Curriculum for the Gifted and Talented

The principles of the cognitive field learning theory are most compatible with the learning traits and characteristics of gifted and talented learners. A curriculum based on the cognitive field theory places emphasis on the whole, the Gestalt. Students are involved in drawing conclusions, looking for evidence as support for conclusions drawn, looking for motives and causes and effects, making comparisons and contrasts, and looking for relationships that are not readily apparent.

When cognitive field theory is applied, gifted and talented students are intrinsically motivated and are involved in analyzing, synthesizing, and evaluating. The teaching methods applied involve the students in open-ended approaches to discussion and debate. Gifted and talented students are viewed as producers of knowledge and engage in problem solving, emphasizing processing that leads to a multiplicity of responses as opposed to arriving at right answers. Students are looking intrinsically for responses rather than concentrating on their notes (which have been provided by the teacher) or looking in the textbook for answers to questions.

The curriculum for the gifted and talented student is characterized by independent student research and complex, in-depth, and advanced organization of content, processing, and products. The content is organized around broad-based issues, problems, questions, or themes. The curriculum addresses ongoing themes and generalizations. The student is studying fractions or reading stories and is also making connections to larger issues of survival, interdependence of humans, and patterns in life, in the environment, or in music. Big-picture issues do not become lost in the struggle to learn important facts and bits and pieces of information. It's not just learning colors in kindergarten but focusing on how color enhances one's understanding of the world and who one is. Time is spent looking at the *color* of a story, or the *color* of a poem, or the *color* represented by a musical piece. Students study issues, problems, and themes that have wide applicability to subjects beyond the most immediate subject. In a unit on discovery in the history class, the teacher does not end the unit before looking at what students have discovered about themselves as they have researched historical discoveries. The organization of the curriculum must cause students to find relationships and make connections with other subjects and other experiences encountered.

The curriculum for gifted and talented students is also characterized by higher-order thinking processes. The emphasis is on processing rather than the product, the right answer. Students are involved in analyzing, synthesizing, evaluating, and problem solving. Students spend time creating novel solutions, thinking and combining the known and the unknown in the creation of something new and different. The creative thinking skills, fluency, flexibility, originality and elaboration are developed and students scamper for additional ideas.

Products in the gifted and talented classroom are real-world products that have real-world applicability. The products are meaningful and can be used in solving problems faced by the school, the community, the state, the nation, or the world. Products are not just completed by students only to be turned in, graded by the teacher, and returned to the student. Products are valued beyond the walls of the classroom and beyond the walls of the school. The teacher finds a way to publicly display or to exhibit products to an audience external to the classroom or school. Students may write solutions to a problem faced by the community and present their thoughts to the local city council for

review. Students may even appear before the school board to present their ideas. The ideas may be published in the editorial section of the local newspaper or mailed to the state legislature. At any rate, there is some effort to present the product to an external public, a real-world audience. A game designed or a play or poem written may be published or presented to an audience at the local library or at the local home for senior citizens. A copy of a book written by a student may be placed in the school library, checked out, and read by other students. Products take on a new, real-world meaning in the gifted classroom.

The curriculum for gifted and talented students is also research oriented. Students engage in experimental, historical, descriptive, correlational research, or case studies. Students always start with a statement of what they hope the results will be (a hypothesis). The scientific method used in any real-world laboratory is evident in the research of gifted students, whether they are in kindergarten, high school, or in science or literature classrooms.

Controversial and provocative questions and issues are characteristics of the curriculum designed for gifted and talented students. Cultural diversity is an ongoing part of the curriculum. Humans' moral reasoning also is recognized in the curriculum. The moral reasoning of historical figures, characters in literature, and people in the news is analyzed in an effort to assist students in the development of their moral reasoning skills.

The application of gifted and talented curricular principles creates modifications in the classroom environment and a new definition of what the classroom can be. The role of the teacher becomes that of a facilitator, a guide, a moderator; the classroom is student centered, and students have opportunities to be self-directed, autonomous learners.

The enrichment triad model, the schoolwide enrichment model, the autonomous learner model, Sternberg's triarchic model, Sydney Parnes's creative problem-solving model, Donald Treffinger's self-directed learning model, the model United Nations program, and Benjamin Bloom and David Krathwohl's taxonomy of educational objectives are examples of curricular models that reflect the principles of the gifted and talented curriculum. Strong programs for the gifted and talented apply a combination of models, capitalizing on the strengths that come with using a variety of models. The best curriculum for the gifted includes in-school and out-of-school programs and

service options for gifted and talented students in homogeneous and heterogeneous settings.

Strategies for Teaching the Gifted and Talented

Strategies for teaching gifted and talented students are compatible with the characteristics of those students. Strategies emphasize the acquisition of concepts and the use of inductive approaches. Rote, repetition, drill, and practice are de-emphasized. Pretesting to uncover what students already know precedes instruction, and students are allowed to accelerate through the curriculum. Instruction and curriculum are differentiated to meet the needs of the gifted and talented.

The three highest levels of thinking—analysis, synthesis, and evaluation—involve non-data-based thinking; students are reading between the lines, drawing inferences and conclusions, arriving at generalizations, and looking for motives and causes. These levels involve thinking of many and varied possibilities, thinking in categories, thinking in detail, and thinking of original ideas. At the evaluation level, students are making judgments and statements of preference while supporting their thoughts by using internal or external standards. Classrooms that utilize these levels of thinking are characterized by the discovery approach, idea sharing, and problem solving. In such classrooms, teachers are facilitators of learning, spending much of their time planning and selecting the curriculum and strategies that will create the type of classroom environment where giftedness will emerge and flourish. Students experience the power of the mind and come to know what it's like to think. Students develop thinking skills that can be used in other classes as well as in situations outside of the school.

There are numerous other thinking skills strategies that are appropriate in helping students develop higher-order thinking skills. Hilda Taba's concept development approach involves productive thinking, including grouping, labeling, and arranging thoughts. Students agree and disagree in the absence of answers and solutions previously provided by the teacher or read from a text. Students come to enjoy playing with ideas and creating unusual ways of looking at real-world issues.

Gifted and talented students also are provided ample opportunities to address questions for which there are numerous possible solutions and responses. Students are able to experience what can happen when real-life problems are solved by thinking in unusual, off-the-beaten-track ways. It was Albert Einstein who developed his general theory of relativity by using fantasy and imagination. Many inventions were developed while inventors toyed with ideas and engaged in thinking which, on the surface, appeared to be totally unrelated to the final product.

Venn diagrams can be used to help students determine similarities and differences among people, places, events, and things. Two-, three-, and four-way Venn diagrams can be used to help students develop spatial thinking skills and to see comparisons and differences previously unseen. The Venn diagram has a place in every classroom. A two-way Venn diagram can be used to illustrate numerals that have straight lines and numerals that have curved lines and to determine how many have both. A three-way Venn can be used to compare and contrast three different characters in a story or novel, or three people in the news. A four-way Venn is a challenge to those who lack experience in looking at how the circles of the Venn come together. A four-way Venn can be intimidating to students who lack experience working with information in this manner.

Teachers can help demystify analogies by helping students build an analogy chain using vocabulary words from science or names of people and places in history. Teachers list words, people, places, things, and content from a subject. Students look for below-the-surface relationships between one word and another word. The best analogies are represented when a relationship is seen between things or people that appear on the surface to be totally unrelated.

Creative thinking skills can be used to foster higher-order thinking skills in the classroom. Students can think fluently, flexibly, originally, or elaboratively. Think of the most ideas; think of many, different ideas. Think of additional ideas. Think of ways to make an idea more interesting. Think of something you have never thought of before.

Frank Williams has 18 different strategies that can be used to stimulate higher-order thinking. Provocative questions are used. Students find traits that are common to all people or think of examples of change or habit.

As students begin to combine thinking at the lower levels with thinking at the higher levels of intellectuality, they begin to experience the joy of using their minds. The classroom becomes more exciting, and students experience the joy that comes when the

learning environment is a mirror of how the gifted and talented student learns.

Socioemotional Needs of the Gifted

Perhaps one of the most neglected areas in gifted and talented education is the social and emotional adjustment of the gifted and talented student. All learners have socioemotional needs that should be addressed. It does appear, however, that with gifted and talented youth and children, the myths that these students will survive without help and that gifted and talented students are superior in all categories, mitigate the urgency generally associated with socioemotional problems among youth. Extreme levels of intellectual, creative, artistic, psychomotor, and psychosocial abilities among the gifted and talented often are accompanied by extreme levels of socioemotional needs. When education neglects the emotional side of learning, gifted and talented children can become anxious, depressed, and alienated.

Gifted and talented learners need to understand their own giftedness, asynchronous development, tendency toward perfectionism, underachievement and dual exceptionality, multipotentiality, personal and social relationships, and depression and anxiety. All of these socioemotional needs can be addressed through a well-planned curriculum of confluence that stresses affective and cognitive learning.

Joyce E. Miller

See also Intelligence and Intellectual Development; Multiple Intelligences; Teaching Strategies; Triarchic Theory of Intelligence

Further Readings

Davis, G., & Rimm, S. (1994). *Education of the gifted and talented.* Boston: Allyn & Bacon.

Ford, D. Y., & Harris, J. J., III. (1998). *Multicultural gifted education.* New York: Teachers College Press.

Gallagher, J. J., & Gallagher, S. A. (1994). *Teaching the gifted child* (4th ed). Boston: Allyn & Bacon.

Maker, C. J., & (1996). *Curriculum development and teaching strategies for gifted learners* (2nd ed.). Austin, TX: PROED.

Marland, S. P. (1971). *Education of the gifted and talented.* Washington, DC: U.S. Government Printing Office.

Olenchek, F. R. (1998). *They say my kid's gifted: Now what?* Washington, DC: National Association for Gifted Children.

Renzulli, J., & Reis, S. (1985). *A schoolwide enrichment model: A comprehensive plan for educational excellence.* Mansfield Center, CT: Creative Learning Press.

Rogers, K. (2002) *Reforming gifted education: Matching the program to the child.* Scottsdale, AZ: Great Potential Press.

Slocumb, P., & Payne, R. (2000). *Removing the mask: Giftedness in poverty.* Highlands, TX: RFT.

Winebrenner, S. (2001). *Teaching gifted kids in the regular classroom.* Minneapolis, MN: Free Spirit.

GOALS

Goals are cognitive representations of future objects that an individual is committed to approach or avoid. Goals are important in motivation and self-regulation, because they focus people's attention on specific possibilities and guide their behavior toward concrete aims. The topic of goals is important to educational psychologists, as they research how general motivational energy is given specific direction in educationally relevant contexts. In this entry, the goal construct is defined in more elaborate fashion, and achievement goals are defined and conceptualized.

Definition of Goal

A more complete understanding of the aforementioned definition of goal may be acquired by considering each of its facets separately.

Cognitive Representation

A goal is a cognitive representation, which means that goals are inferred, not observable, constructs. Goals are restricted to organisms that utilize a mental apparatus in the regulation process; the actions of plant life are not seen as goal-directed behavior, because the organism is reacting reflexively and not with regard to an envisioned possibility. In the same fashion, physiological functions in humans, such as digestion or blood circulation, are not considered goal directed, because these functions are carried out without any mental imaging of the future.

Future

In goal-directed behavior, a future image is used as a guide to present behavior. Behavior is not just an

immediate, unmediated response to a present stimulus, but it involves prevision of a future possibility that has a *causal* influence on present behavior.

Object

The object of a goal is the focal point of regulation. The term *object* is used broadly to refer to an entity, event, experience, or characteristic that is the central focus of the goal. These objects may take on many different forms and contents; they may be concrete or abstract, physical or psychological, observable or unobservable. The object of the goal is not the goal itself. A goal contains both an object and some sort of commitment regarding that object. For example, in the goal "Keep in touch with my parents," "in touch with my parents" is the object, and "keep" is the commitment. An object and a commitment are both necessary components of the goal construct.

Commitment

Goals are adopted when an individual commits to an action of some sort with respect to a cognitively represented future object. Goal commitment is not an all-or-none process; persons may make various levels of commitment to goals and may adopt multiple goals with respect to the same object. Goal commitment must begin with a conscious intention, but over time, such commitments may become automatized and guide behavior without intention or awareness.

Approach or Avoidance

Approach or avoidance is a fundamental distinction that is applicable to any and all types of goals. Goals are focused either on approach or avoidance. Approach goals involve the pursuit or maintenance of a positively valenced object, whereas avoidance goals involve distancing or staying away from a negatively valenced object. Both approach and avoidance "movement" may involve either physical activity (e.g., "Eat one serving of vegetables each day") or psychological activity (e.g., "Stop hating vegetables so much").

Achievement Goals

The goals most relevant to educational psychology are achievement goals that focus on competence. Such goals may be differentiated in terms of two basic aspects of competence: how it is defined and how it is valenced.

Competence is *defined* by the standard that is used to evaluate it. Three standards may be identified: an absolute or task-based standard, an intrapersonal standard based in one's own past or maximum potential attainment, and an interpersonal or norm-based standard. Absolute and intrapersonal standards are typically collapsed together and called *mastery goals*, and normative standards are called *performance goals.*

Competence is *valenced* in that it is either focused on a positive possibility that one would like to approach (success) or a negative possibility that one would like to avoid (failure). Combining the definition and valence aspects of competence results in four basic types of achievement goals. *Mastery-approach goals* represent striving to approach absolute or intrapersonal competence, for example, striving to do better than one has done before. *Mastery-avoidance goals* represent striving to avoid absolute or intrapersonal incompetence, for example, striving to avoid doing worse than one has done before. *Performance-approach goals* represent striving to approach interpersonal competence, for example, striving to do better than other students. *Performance-avoidance goals* represent striving to avoid interpersonal incompetence, for example, striving to avoid doing worse than other students. These four goals are presumed to comprehensively cover the range of competence-based strivings that students pursue in educational settings.

These achievement goals are thought to have an important influence on the way students perceive, experience, and perform in achievement situations. Generally speaking, mastery-approach and performance-approach goals are viewed as adaptive goals that lead to different types of positive processes and outcomes (e.g., mastery-approach goals are thought to facilitate creativity and intrinsic motivation; performance-approach goals are thought to facilitate performance attainment). Mastery-avoidance and, especially, performance-avoidance goals, on the other hand, are viewed as maladaptive goals that lead to negative processes and outcomes, such as selecting easy instead of moderately difficult tasks, quitting when difficulty is encountered, and poor performance attainment. A considerable amount of research over the past two decades has documented the importance of achievement goals for students' behavior in the classroom.

Achievement goals emerge from many different sources, including internal factors, such as dispositional

motives, and external factors, such as classroom or schoolwide characteristics. Those who study achievement goals are in agreement that teachers, administrators, and parents should do all they can to facilitate students' mastery-approach goal adoption and to discourage students' performance-avoidance goal adoption.

Andrew J. Elliot

See also Motivation

Further Readings

Dweck, C. S., & Leggett, E. L. (1988). A social cognitive approach to motivation and personality. *Psychological Review, 95,* 256–273.

Elliot, A. J. (2005). A conceptual history of the achievement goal construct. In A. Elliot & C. Dweck (Eds.), *Handbook of competence and motivation* (pp. 52–72). New York: Guilford Press.

Elliot, A. J., & Fryer, J. W. (2007). The goal concept in psychology. In J. Shah & W. Gardner (Eds.), *Handbook of motivational science.* New York: Guilford Press.

Pervin, L. A. (Ed.). (1989). *Goal concepts in personality and social psychology.* Hillsdale, NJ: Lawrence Erlbaum.

Grade-Equivalent Scores

Grade-equivalent scores are scores reported on norm-referenced assessments indicating the level at which an individual student performed compared with the average performance of students at other grade levels. The score is on a continuous developmental scale used for describing developmental level and measuring growth. Grade-equivalent scores are reported as a decimal number indicating the performance of a student in terms of grade level and month. For example, a score of 6.4 indicates performance at the sixth-grade level, the fourth month of school. Educational psychologists are interested in grade-equivalent scores for students to help provide a frame of reference for their developmental growth. It is important to understand the method for the development of the score and how the score should be interpreted.

Test publishers generally conduct norming studies twice a year: in the fall and in the spring. Tests are administered to students in various grades. The norms are computed by finding the mean raw scores of students in each grade. The score is based on typical performances of students across grade levels. Therefore, the information is aggregated without regard to individual differences within the group. For example, a test publisher may develop an achievement test for third-grade mathematics. To assign norms, the test is administered to large groups of students in successive grade levels. The raw scores are computed for each student in each grade level. The mean raw score is then assigned as the norm for that grade level at that specific time of year. For instance, if the mean number of questions answered correctly is 27 for a group of students in the fifth month of third grade, then any student who also answers 27 questions correctly will receive a grade-equivalent score of 3.5 (third grade, fifth month).

As test publishers typically only do norming studies twice a year, the conversions for months in between the studies are found through interpolation. Interpolation involves plotting the scores on a bivariate axis. The baseline is divided into 10 parts. For third grade, the units would be 3.0, 3.1, 3.2, and so on, until 3.9. The scores are then plotted on the *y* axis. Once the computed scores are graphed, the points are connected, or *interpolated*, with a straight line and then extended, or *extrapolated*, with a line going beyond the specific grades tested to indicate how students in other grades may also perform. Interpolation and extrapolation are necessary to compute norms because the specific test was not given to all grades, nor was it given at all the different points on the continuum within a grade.

Grade-equivalent scores, while intuitively appealing, are the most misinterpreted scores provided by testing companies. The scores are frequently misused to assess student learning. For instance, if a third-grade student earns a grade-equivalent score of 5.7 on a math assessment, this is not an indication that the student is capable of doing fifth-grade math. Instead, it means that this student performed the same as a fifth-grade student did when taking the third-grade math test. Or, the third-grade student can do third-grade math as well as a fifth-grade student in the seventh month of school can do third-grade math. Additionally, does this mean the student should be promoted to fifth-grade math class? No. It is unlikely that the student has been exposed to, or mastered, fifth-grade math material. In fact, if a third-grade student were given a test designed to assess fifth-grade math, it is unlikely that the student would attain the grade-equivalent score of 5.7.

Not only do the scores convey misinformation, but they also lack sensitivity. A difference of 2 or 3 raw

score points in the lower grades can represent as many as 5 or 6 months achievement. However, in the upper grades, 2 or 3 raw score points may not change the grade-equivalent score at all. This is true because students in lower grades tend to perform more similarly than those in the upper grades, where variability is increased. Additionally, standard deviations of grade-equivalents is not provided and not usually known and, therefore, not comparable, which can lead to misinterpretation of academic achievement.

When used for diagnostic purposes, grade-equivalent scores can be problematic. For example, a student may score at the same percentile on an achievement test for successive years, indicating even academic achievement. However, compared with the norming group, the student's grade-equivalent score may be decreasing. A misinterpretation of this score would be that the student is falling behind his or her peers when this would not be the case. Often this is the cause for the misdiagnosis of learning disorders, causing an overdiagnosis in the upper grades and underdiagnosis in the lower grades. In turn, this may also cause a student who consistently performs one standard deviation above the mean to appear to perform substantially better at each successive grade, even though the performance is consistent, and to be placed in advanced classes for which he or she may not be prepared.

The reliability of grade-equivalent scores is limited, as it is simply a score based on the mean raw score for a particular grade level. The failure to recognize that these scores are simply averages can cause unrealistic expectations for students. Because grade-equivalent scores are frequently misinterpreted, they are most suited to be used as an estimate of a student's developmental level or year-to-year growth. These scores are not appropriate as use for a prescription for grade placement.

Lori Jackson

See also Assessment; Norm-Referenced Tests; Percentile Rank; *T* Scores

Further Readings

Burns, E. (2001). The use and interpretation of standardized grade equivalents. *Journal of learning disabilities*, *15*, 17–18.

Campbell, N. J. (1994). Interpreting scores from standardized tests. *Clearing House*, *67*, 314–316.

Hoover, H. D. (1984). The most appropriate scores for measuring educational development in the elementary schools: GE's. *Educational measurement: Issues and practices*, *3*, 8–14.

GRADE RETENTION

Grade retention, a practice that involves students repeating the same curriculum because of failure to master the required skills, is a practice that has been considerably researched. Rarely does an academic field have such overwhelming evidence against a practice that is ignored, almost universally, in American schools. Grade retention is the exemplar of that phenomenon. Despite the research that, over decades, has been systematically analyzed and that suggests that grade retention is not an effective practice, children and adolescents are likely to continue to be affected by grade retention. As Shane Jimerson and his colleagues note, past syntheses are controversial because the original studies are typically based on pre- and posttest designs of retained students only and not on comparisons between retained and promoted but low-performing students. Jimerson and colleagues further describe a literature that is not investigated in longitudinal fashion and fails to account for initial adjustment after the retention experience.

Thomas Holmes's 1989 synthesis of the retention literature indicated that students who were retained had lower academic achievement and more social and emotional maladjustment (i.e., social adjustment, emotional adjustment, and behavior). Any positive trends made in Holmes's analysis suggested that the results of those retention intervention studies diminished over time as students moved to higher grade-level placements. The most recent meta-analysis, conducted by Jimerson in 2001, supports the previous syntheses and provides further evidence that something other than grade retention is necessary for students who are not academically or socially successful in schools. Jimerson's updated synthesis found that retained students' academic achievement was .39 standard deviations below those who were promoted on time. Regarding students' socioemotional adjustment, retained students were rated below their peers on every social, emotional, behavioral, self-concept, adjustment, and attendance measure.

Despite the harmful effects of retention, it is estimated that 5% to 10% of students are retained

annually, and it is likely that grade retention will continue to be prevalent and may even increase. Many school districts and even state educational agencies have used retention as a high-stakes consequence for students who fail to make benchmarks set for a given grade level. As these schools attempt to comply with the current No Child Left Behind legislation, in an environment where economic resources are not sufficient to ameliorate the learning and behavioral challenges that students have before retention may be necessary, it is likely that the numbers of students for whom this practice is employed will increase. The following review summarizes the major areas in the grade retention literature with a focus on younger children in primary grades, grades in which retention is most likely to occur.

Retention in Primary Grades

Each year, millions of children and adolescents enter American classrooms. Some children and parents may be filled with anticipation for an exciting adventure, whereas others may be filled with great anxiety. Schooling young children in the United States is currently trending toward beginning at ages 3 and 4, with the emergence of universal and prekindergarten programs for children at risk. Kindergarten's programming can take the form of half-day or full-day programs. Regardless of the age of entry or the length of the school day, kindergartens often serve as a child's first experience with formal schooling. Although almost all children have successful experiences during kindergarten, some children experience failure. Confusing for the child and often shocking for the parents, children who are not successful in kindergarten are often presented with the possibility of being held back in kindergarten and facing another year of the same kindergarten curriculum.

In 2000, the National Center for Education Statistics and other researchers report that the numbers of children who are retained in kindergarten have remained relatively constant during the 1990s and into the 21st century. In 1993, 6% of children were retained after their kindergarten year; this percentage declined slightly in 1995 to 5%. In 1998, 5% of all children in kindergarten were in their second year of kindergarten. In addition, 6% of children, who were a year older for what would be considered typical entrance age for kindergarten, enrolled in kindergarten for the first time.

Effects of Grade Retention

Research on grade retention of younger children suggests overwhelmingly that the practice of grade retention does more harm than good for the children who experience it. Although there have been a few studies that have found positive effects of retention for young children, the vast majority of literature warns against this practice. Data from the Early Childhood Longitudinal Survey-Kindergarten Class of 1998–1999 (ECLS-K) demonstrate a lack of effectiveness of retention policies on a national scale. The national data on the academic performance of first- and second-grade children who had been retained in kindergarten indicated that they had lower reading and mathematics knowledge than their on-grade-level peers. Children who were retained also were more likely to receive negative feedback from their teachers.

Other research studies, such as work by Arthur Reynolds and colleagues, revealed that grade retention has negative effects on subsequent cognitive achievement in reading and mathematics at ages 14 and 19. Jimerson also found that students who were retained were more likely to drop out of school by age 19, were less likely to receive a diploma by age 20, were less likely to enroll in postsecondary education programs, received lower education/employment status ratings, were paid less per hour, and received poor employer competence ratings.

Characteristics of Retained Students

In addition to studies that document the negative effects of retention on children, there are also studies that document the characteristics of children who are likely to be identified for retention. Panayota Mantzicopoulos suggested factors such as low levels of parental school involvement and parent's satisfaction with school predicted retention. ECLS-K data indicated that in comparing children who were successful in kindergarten with children who were retained in kindergarten, children who were retained were less likely to have attended preschool the year before kindergarten (63%), were more likely to be male (66% vs. 49%), were more likely to have diagnosed developmental difficulties by the end of first grade (22% vs. 9%), were more likely to live in a family considered to be in poverty (19% vs. 11%), and had parents with less than a high school education (17% vs. 7%). Studies that demonstrate that there are some initial

gains in the year after the retention experience also note that these results fade in either the year after the retention or by third grade. Some of the most promising investigations regarding what happens to the academic achievement of students who have been retained have been conducted by a consortium of researchers associated with the Chicago Public Schools. The data from the Chicago Public School System, which banned social promotion in the mid-1990s and since that time has been retaining between 7,000 to 10,000 students per year in third, sixth, and eighth grades, indicates that these retained students struggle with meeting promotional standards and close to 20% of them were placed in special education within 2 years of the retention decision.

Psychological Issues

With the escalating demands of requirements in kindergarten, the careful preparation of children in meeting their academic and psychological needs is very important. Long-term studies suggest even more harmful psychological issues with being retained during the first few years of school. For example, Angela Fiske and Stacey Neuharth-Pritchett found that children who were retained demonstrated higher levels of depression than did their peers who were low-performing or on grade level once they reached middle school. Melissa Roderick's 1994 study produced a similar result: In that study, children who were retained experienced substantial school disengagement during middle school where nearly one quarter of those students dropped out of school and those who remained experienced significant declines in attendance. Grade retention is the single most powerful predictor of dropping out of school. Children who are retained also demonstrate significantly lower emotional health.

Interventions

Although these studies paint a bleak future for young children, there have been a number of advocacy efforts proposed to minimize early school failure. Early care and education programs, such as preschool, are one of the main interventions advocated. Further work by Reynolds and colleagues suggested that preschool participation at ages 3 and 4 was significantly associated with higher reading achievement, higher math achievement, and lower incidence of grade retention in elementary school. Not all families across the United

States can afford high-quality preschool programs for their children. Local, state, and federal governments have joined in the school readiness efforts by providing high-quality preschool programs such as Head Start. Such early childhood education programs are designed to minimize associations between retention and children from lower-income households, those whose mothers have lower levels of education, and those who come from single-parent households.

Access to preschool has contributed to some children experiencing more success in formal schooling. Other interventions have also been proposed, including one-on-one tutoring, multiage classrooms, transition classrooms (classrooms between kindergarten and first grade), and summer remediation programs. One additional strategy, which has been studied extensively over the past few years, is delayed entry into kindergarten, which is also known as *redshirting* or *holding out.* Children with birthdays that are close to the school cutoff for qualifying for kindergarten, or children whose parents perceive that they are socially immature or physically small for their age, are often the children who are redshirted. Studies suggest that more males than females delay kindergarten entry as well as children who have late summer birthdays. A 2000 study by Beth Graue and James DiPerna noted that redshirts are more likely to receive special education services and be boys with birthdays close to the cutoff date. Delaying kindergarten entry does not create any long-term advantages for children. The research-based conclusion is that holding children out of kindergarten for an extra year is not a beneficial practice.

When children are overage for grade, they and their peers are very aware of this difference. The difference becomes even more pronounced as children move from elementary to secondary education environments. Gary Ladd and fellow researchers suggested that as children enter kindergarten, the degree to which they like school may be significantly related to their subsequent achievement. Ladd and colleagues further noted that liking for school and early academic achievements are distinct from other demographic factors (e.g., family background, formative experience, and readiness for school). D. Byrnes and K. Y. Yamamoto's 2001 study revealed that children themselves have concerns about the reactions of their peers to their status as "school failures." A child's academic self-image and peer popularity, when both are positive, reduced the risk of being retained. Assessment of social skills may also be an important contribution in evaluating readiness

(e.g., cooperation, self-control). Poor behaviors and peer relationships may inhibit academic success.

Parental Concerns

Children's and teachers' concerns about success in school differ from parental concerns. Parents' perceptions suggest that parents do not seem as concerned about the academic rigors of kindergarten. Parents and teachers agree that children must be healthy and socially competent and able to comply with teacher authority (this was rated higher by parents than teachers). Parents also rated readiness as more important than did teachers. Parents believed that it was important for a child to communicate in English and to have basic knowledge and skills. Parents who rate their child's ability to do school work as strong had children who were less likely to be retained. African American and Latina/o parents were more likely than White parents to express concern for their child's readiness for kindergarten. White parents were significantly more likely than other parents to delay sending their children to kindergarten until they were older. In 1997, Jimerson and his colleagues found that children who were retained were more likely to have parents who were less involved in their children's schooling.

Teacher Beliefs

Teachers hold a variety of different beliefs about children's kindergarten success. In one study, when academic achievement was held constant, teachers were more likely to recommend retention for children who were perceived as dependent or socially immature. They also were more likely to recommend for retention boys or children who were young. Teachers who themselves were older were more likely to recommend retention. Teachers who rate children with lower levels of personal maturity are more likely to recommend retention. Lorrie Smith and Mary Lee Shepard suggested that teachers see retention as a solution to unreadiness or incompetence. Incompetence might include developmental immaturity.

In Ellen Tomchin and James Impara's 1992 classic study, they reported that teachers saw retention as an acceptable school practice that prevents students from facing daily failure and that motivates them to work harder. Teachers considered academic performance, maturity, ability, gender, and age as factors in their decisions to retain children. Teachers perceived that retention in younger grades was better than in older grades. Teachers in elementary schools believed that retention was an acceptable practice. Students' academic performance was the most influential factor in retention decisions. In Tomchin and Impara's study, teachers were not familiar with the effects and outcomes of retention.

Reasons teachers gave for retaining students include immature attention span (e.g., listening), social skills, developmental delays (e.g., not ready, motor skills, speech-general), home factors (e.g., absenteeism, family situation, no home support), exceptionality (e.g., specific disabilities such as attention deficit disorder, learning disabilities, behavioral disorders), age (e.g., young, late birthday), academic skills (e.g., letters, numbers, readiness skills, has not mastered academics), physical health (e.g., illness, injury), and other issues (e.g., English as a second language, just starting school, changes school often).

Future Directions

Although the literature on grade retention provides strong evidence that the practice is not successful, there are still major areas associated with the practice that need to be explored. Longitudinal research that focuses on academic as well as social and emotional outcomes is needed. Early research on interventions that would minimize grade retention (e.g., early intervention, alternate schooling arrangements) is beginning to show some evidence that grade retention could be minimized in American schools. Research on how parents are involved in the grade retention decision and how classroom environments or teaching practices may contribute to grade retention has yet to be fully explored. What is clear is that the literature demonstrates that the practice of grade retention is one that should be avoided. The focus for researchers should not be whether or not to retain but what can be done in place of retention to meet the individual needs of children and adolescents.

Stacey Neuharth-Pritchett

See also Assessment; Peer Influences; Teaching Strategies

Further Readings

Byrnes, D. A., & Yamamoto, K. (2001). Academic retention of elementary pupils: An inside look. *Education*, *106*(2), 208–214.

Fiske, A., & Neuharth-Pritchett, S. (2006, March). *Coping with depression in adolescence: Transitions from the middle school years.* Paper presented at the biennial meeting of the Society for Research on Adolescence, San Francisco, CA.

Graue, M. E., & DiPerna, J. (2000). Redshirting and academic retention: Who gets the "Gift of Time" and what are its outcomes? *American Educational Research Journal, 37*(2), 509–534.

Holmes, C. T. (1989). Grade level retention effects: A meta-analysis of research studies. In L. A. Shepard & M. L. Smith (Eds.), *Flunking grades: Research and policies on retention.* London: Falmer Press.

Jimerson, S. R. (2001). Meta-analysis of grade retention research: Implications for practice in the 21st century. *School Psychology Review, 30*(3), 420–437.

Mantzicopoulos, P. (2003). Flunking kindergarten after Head Start: An inquiry into the contribution of contextual and individual variables. *Journal of Educational Psychology, 95*(2), 268–278.

Piotrkowski, C. S., Botsko, M., & Matthews, E. (2000). Parents' and teachers' beliefs about children's school readiness in a high-need community. *Early Childhood Research Quarterly, 15*(4), 537–558.

Reynolds, A. J. (1992). Grade retention and school adjustment: An explanatory analysis. *Educational Evaluation and Policy Analysis, 14*(2), 101–121.

Roderick, M. (1994). Grade retention and school dropout: Investigating the association. *American Educational Research Journal, 31*(4), 729–759.

Tomchin, E. M., & Impara, J. C. (1992). Unraveling teachers' beliefs about grade retention. *American Educational Research Journal, 29*(1), 199–223.

GRADING

Grading is a process in which a teacher constructs context, meaning, format, and expected learning outcomes for a unit of academic instruction; selects protocols (multiple-choice items, essay questions, term papers, projects, performances, etc.) to determine how well the student has achieved those outcomes; and then reports the "grade" to the student and/or official agencies. To an outsider, grading may appear as perfunctory as writing a few marginal comments on a C+ essay or scoring an objective test with an 85%, but to experienced educators, grading is clearly a dynamic and demanding classroom interaction.

Historically, teachers have complained about grading, students have complained about grades, and institutions have tinkered with grade reporting methodologies (whole letter grades, plus and minus grades, pass and fail grades, etc.); nevertheless, grading is still here. It is the most entrenched and enduring educational practice found in American secondary schools, community colleges, colleges, and universities.

Viewpoints on Grades

Faculties tend to be very homogeneous in their view of what grades (the isolated artifacts A-, D, or B+ given at the completion of an assignment or course) represent. Faculty members acknowledge that grading criteria and methods of grade calculations may differ among their colleagues, but they all view grades as constructs that recognize and reward various levels of achievement. On the other hand, student life studies suggest that early career students commonly perceive grades as a reward for effort. Effort is important, but it does not directly correlate with academic success. Realistically, students begin classes with different entry-level skills and unique scholastic aptitudes. In any given course, some students will have to put forth a huge amount of sustained effort just to make a C, whereas others will only have to exert a minimal amount of concerted effort to earn a B. Naive views of the role of effort in a competency-based grading system may explain why many students with a 57% cumulative test and assignment average doggedly persist in believing that they will surely get a C in the course if they come to class "almost every time." Obviously, it is imperative that teachers communicate with students early in the academic term about the nature and structure of the grading process.

Grading in a New Paradigm

Intuitively, teachers begin shaping a course by asking themselves, "What am I going to cover in class this semester?" Barbara Walvoord and Virginia Anderson, who have led over 300 grading and assessment workshops in the United States and abroad since 1995, encourage faculty to start by asking a very different question: "What will my students be able to know and do at the end of the semester?" Their student-centered approach to planning and grading was indicative of the major educational paradigm shift of the 1990s—a shift from an emphasis on teaching to an emphasis on student learning. U.S. schools, colleges, and universities had long been evaluated by how well-qualified they were to deliver an educational product—for example, number of state-certified teachers in the

system, percentage of faculty holding a Ph.D., or annual number of scholarly articles published per institution. But by the late 1990s, national education shareholders were asking for direct evidence of student learning—for example, student licensure exam pass rates, writing portfolios, employer ratings of recent graduate skills—to document that students were learning.

Linking Grading and Learning: The Syllabus

The shift from teaching to student learning is best reflected in the course syllabus. What was once a simple "opening day" handout teachers used to explain what topics they would cover, what assignments they required, when tests would be given, and how grades would be averaged has become a mandatory, student-oriented document. Constructing the syllabus is the first step in effective grading. It is also the most critical step in communicating with students about course expectations and grades. Syllabi should list major student learning goals, explain how students can work toward achieving those goals, and explain how their progress will be graded. For example, a 300-level biology teaching course syllabus might begin as follows:

> Welcome to Life Sciences 303! When you successfully complete this course, you will be able to (1) recall and utilize selected vocabulary and information about plants, animals, and ecology; (2) plan and engage children in hands-on science discovery and inquiry activities; (3) demonstrate science safety skills and ethical treatment of animals in the K–8 classroom; (4) select and employ K–8 instructional technologies; and (5) look forward to being a competent and confident elementary science teacher!

The syllabus should convey excitement for the academic discipline and stress the "greater expectations" for student learning (a 21st-century goal stressed by the Association of American Colleges and Universities). The syllabus is also the place to inform students of the consequences of academic dishonesty. Finally, teachers should identify in writing what artifacts (tests, papers, group projects, etc.) will be graded and how the graded works will be counted for the course grade. Exacting attention to grading issues on the syllabus is essential, because

the average student may have to "navigate" five grading schemas a semester.

Grading Strategies for an Objective Test

Scoring a 50-item, multiple-choice test is easy because, as Linda Suskie explains in Assessing Student Learning, objective test items require no professional judgment to score. Grading that same test requires many professional decisions. If a teacher decides to use a criterion-referenced grading system and the student gets 43 of 50 items correct, the student will receive a score of 86%; the teacher may use a conversion table on the syllabus to convert it to a B. If the teacher uses a norm-referenced grading system, that is, "grades on a curve," the student will get a letter grade determined by where 43 fell on the class distribution curve. If 44 were the top test score, the student would receive an A.

Grading Strategies for an Essay Test

Scoring an essay examination requires professional judgment. Despite the time commitment, some teachers prefer grading essay exams because it gives them a better picture of how well their students are learning content, using critical thinking skills, and engaging in significant learning experiences, as operationally described by L. Dee Fink.

Holistic Grading

Many teachers grade essays holistically; this means they assign one letter grade to the entire paper. Holistically graded tests may contain marginal notes (snippets of immediate feedback) or more structured endnotes that offer "big picture" comments.

Grading With Rubrics

Faculty often develop some type of rubric (scoring guide) for grading essays or written assignments. The simplest rubric is a checklist with important features such as "Introduction present" and a place for marking yes or no. Rating scale rubrics add broad categories to the checklist so that the grader can check "Good, Fair, Poor."

As the teacher becomes more descriptive with the rubric, the tool becomes more powerful. Here is

a descriptive rubric that could be helpful in grading any "compare and contrast in 4 ways" essay question, regardless of discipline. The rubric gives a 4-tier scale for grading 3 of such an essay.

Four Comparisons

4 Selected four major comparisons and used key terms

3 Four comparisons, but some superficial or "fuzzy"

2 Three comparisons

1 Inadequate comparison

Four Contrasts

4 Selected four major contrasts and used key terms

3 Four contrasts, but some superficial or "fuzzy"

2 Three contrasts

1 Inadequate contrast

Topic Sentence(s)

4 Each topic sentence is effective

3 Topic sentence(s) present, but misleading

2 Paragraph form evident, but lacks topic sentence(s)

1 Not in paragraph format

To provide maximum feedback to the student, the teacher can circle the performance level on the rubric, calculate the weighted score and grade, and return it with the essay.

Primary Trait Analysis Scales

The real "power tool" in the grading toolbox is the Primary Trait Analysis (PTA) scale. It is a highly explicit, criterion-referenced tool that, as Walvoord and Anderson have demonstrated, has the power to link classroom grading to departmental, programmatic, and institutional assessment. PTA scales work best to assess integrated learning tasks such as making an effective sales PowerPoint presentation, preparing a law brief, or using graphics.

First, the teacher identifies the key "traits" used to understand a graphic in a textbook or newspaper as (a) reading graphic information, (b) using quantitative reasoning, (c) interpreting graphic information, and (d) making connections between graphic information and content information. Then the teacher assigns each trait the same number of criteria. Here is a PTA scale for one of those traits:

Reading Graphic Information

4 Identifies points, ranges, and can verbally describe the function of the graph

3 Cannot verbally describe, but identifies points and ranges

2 Identifies points or ranges

1 Lacks competency

On an essay test, students examine a line graph on a topic and describe fully what they got out of the graphic; on an objective test, they answer questions that have been clustered for the four traits being tested (a strategy called *test-blueprinting*, described by Linda Suskie). The teacher uses the individual student's PTA scores to calculate the student's grade and the PTA scale class averages to inform the teaching process. For example, if the class PTA score for reading graphic information is 3.5 and only 2.1 for using quantitative reasoning, the teacher will engage students in more quantitative reasoning skills this semester and focus on new teaching strategies before next semester.

Advantages of Rubrics

When teachers use descriptive rubrics or PTA scales, they can communicate more effectively with students about grades and grading processes. Because they can also grade papers much more quickly, many teachers now use face-to-face grading conferences with students to establish excellent junior–senior colleague rapport. Student life research indicates students learn best with prompt and specific feedback. Although accrediting national agencies have explained that course grades *cannot* be used for assessment purposes because they reduce all activities to a single letter grade, descriptive rubrics and PTA scales provide specific goal-oriented course-embedded data and are being widely used for assessment purposes.

Strategies for Calculating Grades

How teachers combine graded work to determine a grade is also a strategy decision. If all work is viewed as equal, then total points will work. If different sets of

activities are different in value, each may get a summative grade and these will be averaged for that section grade (e.g., tests count 60%, papers 20%, and quizzes and participation 20%). Definitional grading systems have two sets of grades that affect each other. For example, English teachers stressing emerging writing skills may require a graded work average of B and 75% of all ungraded work turned in order for the student to receive a B; if 70%, then a C grade. Scientists might require both a 90% test average and B lab reports total or an A.

Contract grading is when student and teacher negotiate what will be expected for each grade. In her book *Tools for Teaching*, Barbara Davis offers many other tips for managing grading.

Last But Not Least

National Study of Student Engagement surveys document that teachers can be "hard" and get good student evaluations or "easy" and get bad evaluations; the bottom line is that teachers have to be fair. Fair is testing and grading "what matters most," not disassociated fact tidbits. Fair is grading frequently, promptly, and impartially. Fair is planning and grading for student and learning not simply covering the material. Effective grading empowers students to actively inform and improve their own learning.

Virginia Johnson Anderson

See also Assessment; Cheating; Testing

Further Readings

Association of American Colleges and Universities. (2002). *Greater expectations: A new vision for learning as a nation goes to college.* Washington, DC: Author. Retrieved June 1, 2007, from http://www.greaterexpectations.org

Davis, B. (1993). *Tools for teaching.* San Francisco: Jossey-Bass.

Fink, L. D. (2003). *Creating significant learning experiences.* San Francisco: Jossey-Bass.

Kuh, G. D. (2001). Assessing what really matters to student learning: Inside the National Survey of Student Engagement. *Change, 33*(3), 10–17, 66.

Suskie, L. (2004). *Assessing student learning: A common sense guide.* Boston, MA: Anker.

Walvoord, B. F., & Anderson, V. J. (1998). *Effective grading: A tool for learning and assessment.* San Francisco: Jossey-Bass.

H

People often say that motivation doesn't last. Well, neither does bathing—that's why we recommend it daily.

—Zig Ziglar

HABITUATION

Habituation is a gradual decline in responding to a constant or repeatedly presented stimulus. Responding usually decreases quickly during early stimulus presentations; the loss of responding is slower during later presentations. Habituation is thought to occur for most if not all stimuli and species of animals, including humans, in both laboratory and natural settings. It is one of the longest known and most extensively studied of the behavioral processes. Habituation probably occurs because constant or repeated stimuli usually do not present either a threat or an opportunity. Therefore, animals that ignore these stimuli can devote more energy and resources to important stimuli (e.g., food, threats) and will have an evolutionary advantage as a result.

This entry describes some examples of habituation. It defines habituation's companion process, *sensitization.* It reviews some of the factors that speed or slow habituation. Finally, it discusses some frequent misconceptions about habituation. Because habituation is ubiquitous, it provides a potential explanation for declines in behavior in educational settings. An understanding of habituation provides suggestions for slowing or preventing these declines. Habituation also provides a tool for studying what has been learned in nonverbal animals (e.g., nonhuman animals, babies).

Examples of Habituation

Common laboratory examples of habituation include declining gill withdrawal to a repeated touch in *Aplysia*, decreases in the startle response to repeated noises in rats, decreases in alarm calls by chaffinch to the repeated sight of an owl, decreases in courtship responses with repeated exposure to a potential mate in insects, decreases in exploration with time in a novel environment by rats, decreases in escape responses with repeated exposure to a shadow in crabs, decreases in aggressive responses with repeated exposure to an intruder in several species of fish (e.g., convict cichlids, Siamese fighting fish), decreases in salivation to repeated tastes or smells of food by humans, and decreases in observing of a static stimulus in humans.

Examples of habituation in natural settings include decreases in responsiveness to repeated aircrafts in birds, decreases in response calls to the repeated songs of conspecifics in birds, decreases in fear with increasing contact with humans in bears, and decreases in the reactions to human observers in chimps and gorillas.

Sensitization

Habituation had been studied for many years before it was acknowledged that responsiveness to a repeated or prolonged stimulus often, but not always, increases

before it decreases. This initial increase in responsiveness to a stimulus is called *sensitization.* Sensitization is now thought to be a companion process to habituation that opposes the effect of habituation in many, but not all, situations. Sensitization may occur not only during the first few presentations of a stimulus but also after the introduction of an unexpected stimulus from another modality (e.g., a light or noise). For example, suppose that an infant's visual fixation on a mobile is declining over time of exposure to the mobile (habituation). A sudden, unexpected, noise may restore fixation on the mobile because it produces sensitization that opposes the effect of habituation.

Empirical Properties of Habituation

Habituation provides a useful tool for understanding behavior because its empirical properties have been extensively studied and are well known. An understanding of these properties can be useful for slowing or preventing habituation in situations in which it is harmful (e.g., to the sound of the teacher's voice) or for speeding habituation in situations in which it is beneficial (e.g., reducing initial fear in a new classroom). Although any one of these empirical characteristics might not appear for a particular stimulus or species, most of the following characteristics should be present in any instance of habituation. The first few properties are described in detail because they are studied often, help to distinguish habituation from other processes, and reveal important information about habituation. Later properties are mentioned only briefly.

Spontaneous Recovery

Responsiveness to a habituated stimulus recovers when that stimulus is absent for a time. For example, the taste of ice cream is salient during the first few spoonfuls of a dish. As you consume more of the ice cream, your experience of the taste declines (habituation). Taste returns when you have your next bowl of ice cream on the next day (spontaneous recovery).

Stimulus Specificity

Habituation is disrupted by changes in the presented stimulus. Suppose that you measure decreases in a rat's startle response to a repeatedly presented tone. The startle response will increase again if you change some aspect of the tone (e.g., increase or decrease its frequency, increase or decrease its loudness).

This characteristic of habituation makes it a useful tool for exploring the conceptual universe of nonverbal animals. Suppose, for example, that you suspend a mobile over an infant's crib. The infant will stare at the mobile when it is first introduced, but the time of visual fixation will gradually decline (habituation). The infant will begin to stare at the mobile again if the mobile is changed in some way (e.g., its color). However, this is true only if the infant registers the change. The color of the mobile may be changed, but if the infant does not register the color change, the infant will not begin to stare again.

Variety Effects

Habituation occurs more slowly for stimuli that are presented in a variable, rather than in a fixed, manner. For example, students quickly will stop hearing (i.e., will habituate to) a teacher's voice if that voice is relatively constant. They will habituate more slowly, or not at all, if the teacher modulates his or her voice (varying its loudness, pacing, location, etc.).

Dishabituation

Presenting a strong, different, or extra stimulus restores responsiveness to a habituated stimulus. Have you ever wondered why many restaurants are noisy? Here's one possible explanation. Suppose that habituation to the taste of food helps to terminate eating. In that case, people may eat more in noisy environments because noise acts as a dishabituator that slows habituation to taste and, therefore, encourages eating.

Sensitization and dishabituation are frequently confused because both may involve a decrease in habituation caused by the introduction of a stimulus from another modality. Results are conventionally described as *dishabituation* if the added stimulus restores responsiveness to an already habituated stimulus and as *sensitization* if the added stimulus increases responding before substantial habituation occurs.

Long-Term Habituation

Long-term habituation is habituation that is learned and persists over time. Once an animal has habituated to a stimulus, the animal may not be as responsive to that stimulus as it was in the past. As a result,

successive habituations to a stimulus may occur faster and faster. Long-term habituation also prevents complete spontaneous recovery. Instead, the size of spontaneous recovery may decrease with each repetition of habituation followed by spontaneous recovery.

Dishabituation Habituates

The repeated presentation of dishabituators reduces their ability to restore habituated responding. That is, the effectiveness of a stimulus as a dishabituator is strongest during its first few presentations. It loses effectiveness with repetition.

Stimulus Rate

Faster rates of stimulus presentation yield faster and more pronounced habituation than slower rates.

Stimulus Rate and Recovery

Spontaneous recovery may be faster after faster, than after slower, rates of stimulus presentation.

Stimulus Exposure

Responsiveness to a repeatedly presented stimulus decreases with increases in stimulus exposure.

Stimulus Intensity

Habituation is sometimes, but not always, faster and more pronounced for less, rather than for more, intense stimuli.

Generality

As indicated earlier, habituation occurs for most, if not all, stimuli and species of animals. The exact rate of habituation depends on the species, the stimulus, the response, and the individual subject.

Misconceptions About Habituation

People hold many misconceptions about habituation. First, it is sometimes assumed that habituation is entirely unlearned. However, habituation has both learned (long-term habituation) and unlearned (short-term habituation) components. Because of its learned component, habituation is often thought to be the simplest or most basic learning process.

Second, some people argue that habituation occurs only for reflexive behavior. That is, habituation is restricted to relatively simple behaviors that are automatically evoked by a stimulus, such as a blink to a puff of air delivered to the eye. If that were so, habituation would be of little interest to educators who are concerned primarily with more complex, voluntary behaviors. It is true that early studies examined the habituation of reflexes, and some researchers concluded that habituation occurs only for reflexive behavior. However, it is now known that habituation occurs for a wide variety of relatively complex behaviors that are not evoked directly by a prior stimulus. Exploration provides a good example. The decrease in exploration that occurs over time in a new location is frequently attributed to habituation.

Third, some people assume that habituation occurs only to relatively neutral stimuli, such as lights and noises, and not to biologically important stimuli, such as food or water. However, recent evidence indicates that animals do habituate to biologically important stimuli and that habituation may play a role in the termination of those behaviors. For example, habituation may occur to potentially dangerous aversive stimuli (e.g., the presence of an owl for chaffinch) and to biologically important stimuli (e.g., the taste of food in humans, the presence of a mate in insects).

Finally, habituation is often confused with other processes that produce a decrease in responding (e.g., sensory adaptation, fatigue, extinction, forgetting). These processes should be distinguished, because the decreases in responding occur for different reasons for the different processes. In the simplest form of sensory adaptation, exhaustion of the neurotransmitter in the sensory system reduces responsiveness to a sensory stimulus. Neurotransmitter exhaustion does not produce habituation. Instead, during habituation, the animal detects but does not respond to the stimulus. For example, studies of stimulus specificity show that changes in the stimulus restore habituated responding. Responding would not recover if the animal did not detect the stimulus because the neurotransmitter was exhausted. Fatigue is a decrease in responding that occurs because of muscular exhaustion. Habituation is not caused by muscular exhaustion. For example, habituation occurs even when the animal is exposed to the stimulus but not allowed to use its muscles to respond. Although it is beyond the scope of this entry, *extinction* refers to a decrease in a classically or operantly conditioned response when the conditions that

maintained that response are removed. Extinction is not synonymous with habituation. For example, habituation occurs to responses that have not been conditioned (e.g., to unconditioned reflexes). Finally, forgetting refers to a decrease in behavior that is caused by a failure of memory. In contrast, stimulus memory is relatively good during habituation. Studies of stimulus specificity show that changes in a habituated stimulus restores responding to that stimulus. Changes in the stimulus could not restore responding if the animal had forgotten the stimulus.

Frances K. McSweeney and Eric S. Murphy

See also Classical Conditioning; Learning; Motivation; Operant Conditioning; Reinforcement

Further Readings

Groves, P. M., & Thompson, R. F. (1970). Habituation: A dual-process theory. *Psychological Review*, *77*, 419–450.

McSweeney, F. K., & Swindell, S. (1999). General-process theories of motivation revisited: The role of habituation. *Psychological Bulletin*, *125*, 437–457.

Raynor, H. A., & Epstein, L. H. (2001). Dietary variety, energy regulation and obesity. *Psychological Bulletin*, *127*, 325–341.

Rovee-Collier, C., & Barr, R. (2001). Infant learning and memory. In G. Bremner & A. Fogel (Eds.), *Blackwell handbook of infant development* (pp. 139–168). Malden, MA: Blackwell.

Swithers, S. E., & Hall, W. G. (1994). Does oral experience terminate ingestion? *Appetite*, *23*, 113–138.

Thompson, R. F., & Spencer, W. A. (1966). Habituation: A model phenomenon for the study of neuronal substrates of behavior. *Psychological Review*, *73*, 16–43.

Wagner, A. R. (1976). Priming in STM: An information processing mechanism for self-generated or retrieval generated depression in performance. In J. J. Tighe & R. N. Leaton (Eds.), *Habituation: Perspectives from child development, animal behavior and neurophysiology* (pp. 95–128). New York: Wiley.

HALO EFFECT

The term *halo effect* refers to the tendency to overgeneralize the evaluation of an individual's positive or negative qualities, such that the individual is regarded as chiefly good or chiefly bad, when in fact not enough information is known to accurately apply a positive or negative label to specific qualities. In short, the halo effect occurs when one assumes that an individual with one or two positive traits must have mostly positive qualities in all areas of functioning. Educational psychologists consider the halo effect a cognitive attribution bias, as it involves the unfounded application of general judgment to specific traits.

Halo Effect in the Classroom

In the classroom, teachers are subject to the halo effect rating error when evaluating their students. For example, a teacher who sees a well-behaved student might tend to assume this student also is bright, diligent, and engaged before that teacher has objectively evaluated the student's capacity in these areas. When these types of halo effects occur, they can affect students' approval ratings in certain areas of functioning and can even affect students' grades.

Two cases in which the halo effect might be particularly troubling for students are the case of the gifted but unengaged student and the case of the student with attention deficit hyperactivity disorder (ADHD). Teachers may overlook giftedness when the gifted child is not challenged and behaves differently in class than the typical model student. Likewise, teachers might misjudge the potential of the student with ADHD who appears to be a "nuisance" in class. The student who is not engaged, not putting forth full effort, or not "behaving" becomes the student who is undervalued by his teacher, as these qualities lean toward a general categorization of the student as primarily bad.

Attractiveness Stereotype

A specific kind of the halo effect is the attractiveness stereotype. The *attractiveness stereotype* refers to the tendency to assign subsequent positive qualities to physically attractive people. Research on this subject has shown that even when they have no evidence on which to base their judgments, people often equate physical attractiveness with social competence, higher morality, better mental health, good behavior, greater work achievements, and more frequent successes.

However, contrary to the halo effect, people also tend to judge attractive individuals as more vain and more dishonest than less-attractive people. According to Alice Eagly and colleagues, two traits for which the attractiveness stereotype does not affect evaluations are

integrity and concern for others. Therefore, although the attractiveness stereotype acts like a halo effect in that observers tend to prejudge attractive people as mostly good, it seems to be less powerful than a typical halo effect because it extends to a few negative qualities as well.

The attractiveness stereotype exists in a wide range of cultures, but the good qualities associated with being attractive vary depending on the specific culture. Douglas Johnson and John Pittenger found evidence of the attractiveness stereotype still being relevant among elderly individuals in Western cultures. However, research examining the generalizability of the attractiveness stereotype to children, and whether the stereotype extends to the "cute" child in the classroom, is scarce.

History of the Halo Effect

The halo effect is prevalent today, but it is not a new social phenomenon. The first landmark study that revealed the existence of the halo effect occurred in 1920, when Edward L. Thorndike studied military superiors' evaluations of officers in training. In this study, superior officers were instructed to rate separately training officers' abilities in the areas of physical qualities, intelligence, leadership, and character. In this task, evaluating officers were directed to completely disregard every trait of the rated individual except the one trait being evaluated in each section. Despite these explicit and precise instructions, for each training officer, raters provided scores that were much more highly correlated than would normally be expected among the four performance categories.

For instance, Thorndike discovered that the intelligence ratings one conscientious supervising officer gave to nearly 140 officers in training correlated .51, .58, and .64 with his ratings of physique, leadership, and character, respectively. Other evaluating officers' ratings showed a similar trend. Thorndike noted that, theoretically, intelligence should relate much more closely to leadership or character than physique and these high, even correlations speak to a constant error in evaluation.

Other evidence for this rating error appeared in evaluations of aviation cadets' flying abilities. Because of the high level of specific technical ability required in flying, Thorndike estimated that correlations between competency ratings in flying and in general officer work should be .40 or lower, and .25 or lower with

attenuation. Instead, he found the average correlation between these two areas for the eight raters was .67, with half of the raters' scores having correlations stronger than .72.

Another early study of the halo effect, completed by F. B. Knight and R. H. Franzen, showed similar surprising results with ratings of teachers. Perhaps the most concerning finding is that raters' scores of teachers' voices correlated relatively strongly with scores of intelligence, interest in community affairs, and general merit as a teacher. Judgments in one area of competency affected judgments in other areas, yet these ratings still determined salaries and promotions. These findings suggest that sometimes the halo effect does not diminish, even when stakes are high and personal.

Future Research

Thorndike was the first person to name this trend a halo of general merit and noted that he believed even the most capable of people are prone to clouded judgment from the halo effect. As a suggestion for minimizing bias, he recommended evaluators cite specific evidence before rating multiple abilities. Since then, some researchers have confirmed the presence of these effects in modern society, whereas others have concluded that halo effects are not powerful enough to have significant consequences. More research in this area would help determine the relevance of halo effects and conclude how best to avoid these effects if they are indeed substantial.

Kristin Rasmussen

See also Evaluation; Grading; Stereotypes

Further Readings

Eagly, A., Ashmore, R., Makhijani, M., & Longo, L. (1991). What is beautiful is good, but . . . A meta-analytic review of research on the physical attractiveness stereotype. *Psychological Bulletin*, *110*, 109–128.

Johnson, D. F., & Pittenger, J. B. (1984). Attribution, the attractiveness stereotype, and the elderly. *Developmental Psychology*, *20*, 1168–1172.

Kaplan, R. (1978). Is beauty talent? Sex interaction in the attractiveness halo effect. *Sex Roles*, *4*, 195–204.

Knight, F. B., & Franzen, R. H. (1922). Pitfalls in rating schemes. *Journal of Educational Psychology*, *13*, 204–213.

Thorndike, E. L. (1920). A constant error in psychological ratings. *Journal of Applied Psychology*, *4*, 25–29.

Wheeler, L., & Kim, Y. (1997). What is beautiful is culturally good: The physical attractiveness stereotype has different content in collectivistic cultures. *Personality and Social Psychology Bulletin*, 23, 795–800.

HEAD START

Created in 1965, *Head Start* is a federally funded child development program designed to promote school readiness by improving the cognitive and social-emotional development of primarily low-income children. This program fosters their development by providing educational, health, nutritional, and social services to these children and their families. Viewed by many Americans as one program, Head Start actually is two programs: *Head Start* and *Early Head Start.* Established during the 1994 Reauthorization of the Head Start Program, Early Head Start is a comprehensive early childhood program serving primarily low-income children prenatal to age 3, pregnant women, and their families. This entry examines the origins of Head Start and Early Head Start, the administration and funding of these programs, program eligibility, demographic characteristics of these programs, the Head Start Program Performance Standards and Other Regulations, parental involvement in these programs, these programs' assessment and monitoring requirements, empirical research evaluations of these programs, and policy considerations.

Origins of Head Start and Early Head Start

In 1964, President Lyndon B. Johnson launched the War on Poverty. To help wage this war, federal lawmakers enacted the Economic Opportunity Act of 1964. This act created the Office of Economic Opportunity (OEO), and Sargent Shriver was appointed as OEO's first director.

Shriver, a brother-in-law of the late president John F. Kennedy and a former president of the Chicago Board of Education, considered an early intervention program for poor children for two reasons. First, Shriver quickly learned, in his new position, that many poor people were children. Second, he was familiar with research, funded by the Joseph P. Kennedy, Jr. Foundation, which found that the IQs of mentally retarded children can be increased through exposure to positive early interventions and that proper nutrition can have a positive impact on a child's intellectual development. If mentally retarded children could benefit by participating in an early intervention program, Shriver believed that poor children could also benefit from participating in an early intervention program.

A multidisciplinary committee of 14 child development experts, including developmental psychologists, was assembled by Shriver to conceive an early intervention program for poor children. Chaired by Robert E. Cooke, professor of pediatrics at Johns Hopkins University School of Medicine and chairman of the Kennedy Foundation's Scientific Advisory Committee, this multidisciplinary committee considered how to create a national early intervention for poor children. In February 1965, this committee recommended the establishment of Project Head Start. This program would provide poor children and their families with a comprehensive range of services, including education, nutrition, physical and mental health services, and social services, and this program would mandate considerable parent involvement.

Substantial parent involvement in the decision making of Head Start programs was influenced by the civil rights movement and the concept of "maximum feasible participation." Under this concept, policymakers believed that the effectiveness of federal programs serving poor people would be enhanced if the poor people were involved in the planning and administration of these programs.

From Shriver and the multidisciplinary committee's work, the federal government launched Project Head Start in the summer of 1965 by providing funds to local community programs. According to Valora Washington and Ura Jean Oyemade, local Head Start organizations throughout the United States provided an 8-week summer program of services to 561,000 children. Head Start's federal budget was $96.4 million. Federal funding for Head Start bypassed state governments because many southern governors were segregationalists and opposed the program.

In the years after 1965, the Head Start program's duration changed, as did Head Start's place in the federal bureaucracy. Head Start ceased to be a national summer program, and many programs became 9-month half-day programs or 9-month full-day programs. In 1969, the federal administration of Head Start shifted from OEO to the Office of Child Development in the then U.S. Department of Health, Education, and

Welfare. The administration of the Head Start program was later moved to the Head Start Bureau in the Administration of Children, Youth, and Families (ACYF) in the Administration for Children and Families (ACF) within the U.S. Department of Health and Human Services (DHHS). In 2006, the Head Start Bureau became the Office of Head Start, and the Office of Head Start was elevated from ACYF into ACF within the DHHS.

In December 1993, about 20 years after federal policymakers created the Head Start Program, a federal Advisory Committee on Head Start Quality and Expansion recommended expanding Head Start services for infants and toddlers. This committee gave three reasons for its recommendation. First, recent biological research revealed that tremendous child development takes place before a child reaches age 3 but that a child during this period of his or her life needs a favorable environment and support from his family and community. Second, negative national trends, such as rising poverty, decreasing access to adequate health care and quality child care, and an increasing rate of teen pregnancies, endangered young children's exposure to a favorable environment for development. Third, research showed that providing early assistance to young children and their families can positively affect the development of young children and families.

Subsequently, federal lawmakers decided to create the Early Head Start program by enacting the Head Start Authorization Act of 1994. The Early Head Start program started to serve infants and toddlers and their families in 1996.

Administration and Funding

Head Start and Early Head Start are federal-to-local programs. The Office of Head Start administers the Head Start and Early Head Start programs at the federal level. ACF, in coordination with the Office of Head Start, awards competitive, multiyear grants to organizations in the local community. These organizations are called grantees. Some grantees do not provide direct services but contract with another organization, called a delegate agency, to provide those services.

According to the Head Start Program Information Report (PIR) data for the 2005–2006 program year, there were 2,696 Head Start and Early Head Start programs. Of these local programs, 38% were operated by private or public nonprofit organizations (non–community action agencies, such as churches or

nonprofit hospitals), 31% of these programs were operated by community action agencies, 17% by school systems, 7% by tribal governments or consortia (American Indian/Alaska Native), 6% by government agencies (non–community action agencies), and 1% by private or public for-profit organizations (e.g., for-profit hospitals).

Because Head Start and Early Head Start are administered as federal-to-local programs, they receive 80% of their funding from the federal government and must match those funds with assistance from the local community. This 20% local match can consist of cash, donations, or volunteer hours. During fiscal year 2007, federal lawmakers appropriated $6.8 billion to fund 1,951 Head Start and 745 Early Head Start programs.

Eligibility

Unfortunately, the current level of federal funding for Head Start and Early Head Start programs is insufficient to serve all eligible children and families. The National Head Start Association estimates that Head Start and Early Head Start programs only serve about 42% and 2% of Head Start and Early Head Start eligible children and families, respectively.

Children and families are eligible to enroll in Head Start and Early Head Start if the family's annual income is below the federal poverty guideline and/or if that family is receiving public assistance, such as the Temporary Assistance to Needy Families program. Head Start and Early Head Start programs must make available at least 10% of their enrollment for children with disabilities and can allow up to 10% of their enrollment to consist of children whose family has an income above the federal low-income guideline.

Demographic Characteristics

According to PIR data, Head Start programs served nearly 1 million children, and Early Head Start programs served 85,831 children and 10,825 pregnant women during the 2005–2006 program year. Of the more than 1 million children that the Head Start and the Early Head Start programs served, 131,695 were children with disabilities.

Head Start and Early Head Start programs vary in terms of race and ethnicity, according to the PIR data. Forty percent of children and pregnant women are White, 31% are Black or African American, 6% are biracial or multiracial, 4% are American Indian or

Alaskan Native, 2% are Asian, 1% are Native Hawaiian or other Pacific Islander, and 16% were reported as other or an unspecified race. In terms of ethnicity, 34% of Head Start and Early Head Start children and pregnant women are Hispanic or of Latina/o origin.

Head Start Program Performance Standards and Other Regulations

The Head Start Program Performances Standards and Other Regulations (hereafter referred to as "standards") set the minimum national standards that Head Start and Early Head Start programs must meet. These standards are revised periodically to reflect best practices and ensure that every program provides a comprehensive and consistent set of quality services. In fact, Head Start programs have been rated as "good"-quality programs using the Early Childhood Environment Rating Scale: Revised Edition. These standards require that each Head Start and Early Head Start program provide nutrition, health, educational, social, and family services and that children and families with disabilities receive appropriate, inclusive, and individualized services. Within the context of the standards and the Head Start Child Outcomes Framework, every program must use curricula and assessments with their children that address each of the eight developmental domains: language development, literacy, mathematics, science, creative arts, approaches to learning, physical health and development, and social and emotional development.

On the other hand, these standards are crafted so that Head Start and Early Head Start programs have much flexibility in how they implement these standards. For example, each local Head Start program is permitted to select and use its curriculum because there is no mandated national curriculum. The curriculum must, however, be consistent with the standards and grounded in solid child development principles.

Head Start and Early Head Start programs are able to provide many of the required comprehensive services by collaborating with public schools and federal, state, and local agencies and through partnerships with community organizations, such as public schools, social service agencies, and businesses. Collaboration and partnerships enable programs to implement the standards flexibly and to find creative solutions to address community needs.

Partnerships are also an important part of the relationship that each program has with children and families. Head Start programs must present parents with opportunities to form family partnership agreements. Families work with program staff to forge these agreements. These agreements outline family goals, responsibilities, and schedules for families to achieve. Through family partnerships, program staff can help families strategize how to move forward in their lives.

Parent Involvement

Because Head Start and Early Head Start are each two-generational programs, the standards require that programs strongly encourage parental involvement as a way to promote their children's education and overall development. Having parent development in these programs is also important because it fosters the educational and professional development of the parents.

Head Start and Early Head Start parents participate in the management of their children's program by serving on policy councils, policy committees and parent committees. Each grantee agency has a policy council, and each delegate agency has a policy committee. Subject to each program's bylaws, the policy council and the policy committee have decision-making roles in each program. Policy council and policy committee members must consist of 51% parents of enrolled children and 49% local community representatives, such as businesspeople and civic leaders. Parents serving on a policy council or a policy committee must be elected or re-elected each year to 1-year terms, with a term limit of three terms. Parent committees exist at the center level and consist entirely of parents of children enrolled in that center.

By serving on a policy council, policy committee, or parent committee, parents gain self-confidence and important workforce skills. Perhaps as important is that children see their parents in a position of responsibility, maybe for the first time, and from that their children gain self-esteem and self-confidence in what they can accomplish when they reach adulthood.

Policy councils and policy committees are required to serve as a liaison with the parent committees, grantee and delegate agency governing bodies, public groups, private organizations, and other groups in the community. Policy councils and policy committees help recruit volunteers and pro bono services from parents and the local community, mobilize community resources to meet community needs, and develop procedures so that programs can address any community concerns about the program.

Policy councils and policy committees help parent committees to encourage parents to participate in the program and to educate parents about their rights, opportunities, and responsibilities in their program. Policy councils, policy committees, and program staff help parent committees to plan, coordinate, and organize parental activities and to make sure that the program budget has funds available to support these activities. The parent committee provides advice to staff in the development and implementation of the program's policies, activities, and services; participates in the planning and implementation of parent and staff activities; and is involved, to some extent, in recruiting and screening program employees.

Meanwhile, each Head Start and Early Head Start program has a governing body, or board of directors. The governing body consists of individuals from the local community, such as attorneys, doctors, educators, and elected officials. Governing bodies are responsible for ensuring that appropriate internal accounting controls exist to safeguard federal funds and that an annual independent audit of the program is conducted.

In partnership with the program's governing body and program staff, policy councils and policy committees determine how they and the governing body will implement shared governance or decision making. Policy councils and policy committees work with the governing body and program staff to develop, review, and make decisions regarding their program's policy and procedures in the following areas: shared governance, funding or grant applications, budget, program planning, program philosophy, makeup of the policy council or the policy committee, recruitment and enrollment of children, annual self-assessment of program's progress, and personnel decisions. In situations where the governing body conflicts with the policy council or the policy committees on an issue, each Head Start and Early Head Start program has procedures in place to resolve an internal dispute.

Assessments and Monitoring

Assessment and monitoring are important features of every Head Start and Early Head Start program. Shortly after enrolling a program, Head Start and Early Head Start children must undergo health and behavioral screenings to identify any developmental concerns and to make sure that appropriate services are provided to address these concerns. Each Head Start program must have a system in place to assess child outcomes at least three times per year.

In addition to this locally based child assessment, each Head Start program must participate in the National Reporting System (NRS), a controversial national assessment with narrow cognitive and social-emotional measures. All kindergarten-eligible 4- and 5-year-old children whose primary language is English or Spanish are assessed near the beginning and end of their program year. The NRS has been criticized by the U.S. Government Accountability Office and hundreds of child development researchers and assessment experts. Major legislation is currently pending in the U.S. Congress to suspend and terminate the NRS.

Head Start and Early Head Start programs participate in a rigorous monitoring system. Every 3 years, the federal government uses the Program Review Instrument for Systems and Monitoring to review the program's compliance with federal law and the standards. If a program is found out-of-compliance, it must take corrective action or risk losing its federal grant.

Empirical Research Findings

Head Start has been known as the nation's research laboratory for early childhood education. Rigorous and methodologically sound evaluations of the Head Start and Early Head Start programs over the past 40 years have generally found that these programs provide benefits to the children and families served.

However, a major evaluation of Head Start early in its history appeared to show that its benefits were limited and short term. Shortly after Head Start began in the 1960s, the Westinghouse Learning Corporation and Ohio University researchers evaluated Head Start and reported that the benefits of Head Start to its children seemed to fade out by the time these children reached third grade. The methodology employed by these researchers, however, has been criticized, and more recent research has provided two possible explanations for these fade-out effects in the Westinghouse and other studies. One explanation, offered by W. Steven Barnett, is that inappropriate comparisons were made between the children in the intervention and comparison groups of the Westinghouse study. These methodologically flawed comparisons made it appear that fade-out effects occurred when, in fact, none existed. Another explanation, presented by Valeria Lee and Susanna Loeb, is that Head Start children frequently attend middle schools of poorer quality than

their more affluent peers. According to this research, it is the low quality of the middle schools that explains why Head Start children make less progress in middle school than their more affluent peers.

Since the Westinghouse study, much more evaluation research has been conducted, and the evidence shows that Head Start has produced short-term and/or long-term educational, economic, health, and law enforcement benefits to the children and families it has served. Barnett conducted a comprehensive review of the effectiveness of early childhood programs and found that Head Start children experience favorable effects on grade repetition, special education, and achievement test scores.

The Head Start Family and Child Experiences Survey, a nationally representative quasi-experimental study, demonstrated that Head Start children make educational progress during their Head Start and kindergarten years. By the spring of their kindergarten year, Head Start graduates' reading assessment scores reached national norms and their general knowledge assessment, early math, and vocabulary scores were close to national norms.

The Head Start Impact Study, a nationally representative, experimental and random assignment study, showed in its initial findings that Head Start had favorable impacts on children's health and cognitive and social-emotional development after these children had 6 to 9 months of Head Start. This study also found that Head Start parents improved their parenting practices with their children more than the comparison group parents did. The final Head Start Impact Study report is expected to be released in late 2007.

A San Bernardino County, California, researcher tracked the progress of more than 600 Head Start graduates into kindergarten. Based on their progress, he projected their future benefits and costs to society. He found that these Head Start graduates would yield nearly $9 in benefits for every $1 dollar invested in these Head Start children. These benefits included increased employment, family stability, and earnings and decreased crime costs, grade repetition, and special education.

In their recent regression-discontinuity study, Jens Ludwig and Douglas Miller found that Head Start reduced mortality rates for children ages 5 to 9 years from 1973 to 1983 from causes that could have been affected by their participation in Head Start, when they were 3 and 4 years old. Their research also suggested that Head Start had favorable impacts on high school graduation and college attendance rates.

Likewise, Eliana Garces and colleagues found that Head Start had positive effects on high school graduation and college attendance rates; they also found a reduction in criminal activity. Garces and colleagues found that Whites who attended Head Start were more likely than their siblings who did not attend Head Start to have graduated from high school or attended college. Their research also found that African Americans who attended Head Start were less likely than their siblings who did not attend Head Start to have been charged with a crime.

Meanwhile, the Early Head Start has had favorable short-term impacts on the children and families served. When Early Head Start was created, federal lawmakers mandated that an Early Head Start Impact Study be conducted. This impact study was a longitudinal, random assignment study planned to measure Early Head Start effectiveness in selected programs. This study found that Early Head Start children at age 3 had experienced larger impacts on their cognitive, language, and social-emotional development than the comparison group children had; in addition, parents of participating children had experienced improved parenting practices and mental health. These parents also showed progress toward self-sufficiency and were more likely to participate in educational activities than the comparison group parents were.

Policy Considerations

The growing body of favorable empirical research on Head Start and Early Head Start and the strength of the Head Start grassroots community as a political force explain why the U.S. Office of Management and Budget and the U.S. Department of Health and Human Services rated Head Start's performance as "moderately effective" under the 2006 Program Assessment Rating Tool (PART). In an earlier PART review, Head Start program had been rated as "Results Not Demonstrated."

Currently, federal lawmakers are considering legislation to reauthorize the Head Start and Early Head Start programs. It is likely that federal lawmakers will preserve the federal-to-local nature of these programs; encourage Head Start's collaboration with other early childhood programs and services in communities; increase teacher degree requirements; raise the Early Head Start, Seasonal/Migrant Head Start, and American Indian/Alaskan Native Head Start set-asides; end the NRS; maintain comprehensive services; and increase

income eligibility from 100% to 130% of the federal poverty guidelines. Strong parent involvement in the program will likely continue, but the parents' decision-making role in program may be reduced to an advisory role.

Benjamin Allen

See also Cognitive Development and School Readiness; Early Child Care and Education; Emotional Development; Family Influences; Literacy; Longitudinal Research; Parenting; Poverty; School Readiness; Social Development

Further Readings

Barnett, W. S. (1998). Long-term effects on cognitive development and school success. In W. S. Barnett & S. S. Boocock (Eds.), *Early care and education for children in poverty* (pp. 11–44). Albany: State University of New York Press.

Carnegie Corporation of New York. (1994). *Starting points: Meeting the needs of our youngest children* (abridged version). New York: Author. Retrieved April 27, 2006, from http://www.carnegie.org/starting_points/startpt1.html

Garces, E., Thomas, D., & Currie, J. (2002). Longer-term effects of Head Start. *American Economic Review, 92*, 999–1012.

Ludwig, J., & Miller, D. (2007). Does Head Start improve children's life chances? Evidence from a regression discontinuity design. *Quarterly Journal of Economics, 122*, 159–208.

U.S. Department of Health and Human Services, Administration on Children and Families, Office of Head Start. (2006, June 15). *Head Start program fact sheet, fiscal year 2006.* Available from http://www.acf.hhs.gov/programs/hsb/about/index.html

Washington, V., & Oyemade, U. (1987). *Project Head Start: Past, present, and future trends in the context of family needs.* New York: Garland.

Zigler, E., & Styfco, S. (Eds.). (2004). *The Head Start debates.* Baltimore: Brookes.

Zigler, E., & Valentine, J. (Eds.). (1997). *Project Head Start: A legacy of the war on poverty* (2nd ed.). Alexandria, VA: National Head Start Association.

Zill, N., & Sorongon, A. (2004, June 28–30*). Children's cognitive gains during Head Start and kindergarten.* Paper presented at the National Head Start Research Conference, Washington, DC.

HIGH-STAKES TESTING

Large-scale, standardized tests of academic achievement have been used to measure student learning and school effectiveness in the United States since the mid-19th century, and their use continues to be widespread. Although many of these tests are intended to serve primarily as tools for monitoring the education system as a whole or for providing information to help teachers adjust their instruction, in some cases, performance on standardized tests is associated with consequences for individual students, teachers, or schools. The term *high-stakes testing* refers to such uses of tests and includes testing for selection or certification (e.g., college admissions tests, professional licensure tests, tests used to assign students to specific courses or course sequences), testing that is used to determine whether students graduate or are promoted to the next grade level, and testing that is part of a formal test-based accountability system, such as those enacted in response to the federal No Child Left Behind (NCLB) legislation, signed into law in 2002. Although the term *high-stakes test* is somewhat misleading, in that a particular test might have stakes only in certain contexts, the term is often used interchangeably with *high-stakes testing*. This entry focuses primarily on tests used for accountability purposes, though many of the findings and implications are relevant to other high-stakes testing contexts. The remainder of this entry provides a brief description of test-based accountability systems, a summary of research on the effects of high-stakes testing, and a discussion of considerations for ensuring appropriate use of high-stakes testing.

Testing in Accountability Systems

The high-stakes testing that is used for educational accountability dominates many of today's policy discussions about testing. Large-scale achievement tests constitute a central component of what is often called a *test-based accountability* system. Test-based accountability systems typically include four key components: (1) standards that describe what students are expected to learn (often called *content standards* or *academic standards*) and the level of proficiency they must demonstrate (often called *performance standards* or *achievement standards*); (2) tests that are intended to measure attainment of the standards; (3) targets for performance on those tests (which may be simple cut scores or more complex combinations of information from multiple tests, such as the Adequate Yearly Progress [AYP] measures required under NCLB); and (4) a series of consequences that could

include a variety of rewards, sanctions, and interventions. The tests used in these systems typically have several purposes, including motivating educators and students to work harder, shaping instruction to make it resemble the standards, and providing both school personnel and the general public with information about student performance to help promote better decision making. Advocates of test-based accountability argue that attaching high stakes to test scores is essential for ensuring they serve these purposes adequately, whereas detractors argue that stakes distort the quality of information and create incentives for student and teachers to engage in counterproductive behaviors. Some of the research findings bearing on this debate are summarized later in this entry. Accountability testing is distinguished from many other high-stakes testing contexts by the fact that the stakes do not necessarily affect the test-takers but instead are often focused on the educators whose performance is being judged on the basis of students' test scores.

The degree to which a test is considered to have high stakes varies in part as a function of the specific ways scores are used to make a decision, as well as the other criteria that are included in that decision. College admissions tests, for example, might be considered to have substantial stakes in some contexts, such as when applying to highly selective schools that place heavy emphasis on those tests, but not in others, such as when applying to less-selective schools or to schools where admissions decisions permit low scores to be offset by other factors. A state accountability test might have no stakes for individual students but high stakes for principals and teachers working in schools that are at risk of not meeting performance targets. Many of the concerns about high-stakes tests today stem from tests' uses for promotion and graduation decisions at the individual student level and their role in evaluating, rewarding, and sanctioning schools, school districts, or teachers. Both of these categories of test use can be thought of as serving an accountability purpose, and both are likely to affect students as well as educators, even when consequences are directed at only one of these groups. And unlike many uses of tests for admissions decisions, in most cases low scores on these accountability tests are sufficient to provoke a negative consequence, regardless of student performance on other measures. There is a growing body of research examining what happens when test-based accountability systems are put in place.

Consequences of High-Stakes Accountability Testing

The use of achievement tests in accountability systems has been shown to influence a variety of outcomes at all levels of the education system. Most of the research that has examined the effects of high-stakes testing addresses outcomes in one of three categories: instructional practice (broadly defined to include actions taken by teachers and administrators to influence the instruction provided by the school), educators' attitudes or motivation, and student achievement.

Anyone examining results from these studies must keep in mind that the responses observed in any specific testing context are undoubtedly related to the specific features of that context, such as the severity of stakes, the length of time the testing program has been in place, and the types of instructional support provided by schools, districts, or states. Although the results of a study conducted in one context cannot be generalized to other testing situations, for most of the outcomes discussed in this entry, multiple studies have produced similar conclusions and, therefore, provide more grounds for generalizing. It is also important to note that most of the studies of accountability testing focused on school-level accountability systems, such as those enacted in response to NCLB. Teacher-level accountability systems, which often take the form of pay for performance, are becoming increasingly prevalent and are in need of research to explore how the effects differ from those observed under school-level systems.

Effects on Instructional Practice

If test-based accountability is to function as an effective tool for improving education, it must ultimately influence the quality of instruction students receive in the classroom. There is broad consensus that curriculum and instruction are influenced by high-stakes testing, but there also is some disagreement in the policy and research communities about the nature of these effects and whether, on balance, they might be judged beneficial or harmful.

Several frameworks have been created for understanding teachers' responses to testing. One of them, developed by Daniel Koretz and colleagues, includes seven categories of responses. Three of these would generally be considered positive outcomes of high-stakes testing: providing more instructional time (such

as through an increase in overall instructional time or a decrease in time spent on noninstructional activities), working harder to cover more material, and working more effectively. One category, cheating, is unambiguously negative and produces distortions in test scores. The remaining three are somewhat ambiguous, with possible positive or negative effects depending on the specific nature of the responses and the context in which they take place. These are reallocating instructional time to focus more on tested material and less on material that is not tested, aligning instruction with the standards on which the test is based, and coaching by focusing on incidental aspects of the test such as specific item formats or styles. Generally speaking, test-score gains that result from the first three (positive) responses probably provide reasonably valid information about increases in student learning. Gains resulting from cheating are clearly misleading. One of the most challenging aspects of evaluating the effects of high-stakes testing stems from the ambiguous nature of test-score gains that result from reallocation, alignment, or coaching.

Some research suggests that high-stakes testing often leads to the positive responses listed in the previous paragraph. In particular, teachers sometimes report that they work harder and make efforts to improve their own practice when faced with a requirement to increase students' test scores. There have also been many documented instances in which teachers and other school personnel have engaged in cheating. The bulk of the research, however, has emphasized the three ambiguous responses described in the previous paragraph. Most studies of teachers' responses to high-stakes testing reveal extensive reallocation and coaching, and efforts to align instruction with standards have become increasingly common in response to state standards-based accountability systems that emphasize the importance of standards for guiding instruction. Common responses reported by teachers include shifting time away from topics that are not included on the test so that they can devote more time to tested content, having students practice with problems that closely resemble test items (e.g., short reading passages followed by a few multiple-choice items), having students take practice tests constructed from released items from previous test administrations, and having students refer to scoring rubrics when constructing written responses throughout the school year. District-wide changes have also been reported, such as an increase in instructional time

allocated to tested subjects and a decrease in time spent on nontested subjects. In some cases these changes might be considered desirable. For example, states and districts have experimented with open-ended or performance-based assessments as a way to promote the teaching of complex problem-solving skills and other skills that are difficult to measure using the multiple-choice format. In some districts where such tests have been introduced, teachers reported an increased emphasis on the desired skills and processes. At the same time, some efforts to engage students in test-like activity might be considered detrimental, particularly if they result in an exclusion of other valued activities or if they lead to excessive focus on a narrow set of problem types.

In addition to the test itself, the methods used to calculate scores and report performance can influence practice. One of the more salient features of NCLB-mandated testing is the requirement that performance be reported according to whether students score above or below the "proficient" level. Under this system, educators may be urged to move students from below to above the proficient level while receiving no credit for performance changes that do not cross this threshold. These systems appear to have led many teachers to focus more of their attention on students performing near the proficient cut score (often referred to as "bubble kids"), perhaps leading to reduced effort expended on behalf of students performing above that level or students performing far below it.

Effects on Educators' Attitudes

There is less research on testing's effects on attitudes than on practices, but there is some evidence that high-stakes testing systems influence how teachers and students view their school experience. In particular, several surveys indicate widespread teacher reports of reduced morale as a result of testing policies. Majorities of teachers also report that the instructional practices they adopted in response to testing were inconsistent with their views of good teaching.

At the same time, there is evidence that high-stakes testing can lead to improved teacher morale if districts or schools provide appropriate support and learning opportunities for teachers. Teachers also have reported that high-stakes testing has, in some cases, helped them focus their instruction more effectively or has provided useful feedback, but more frequently, teachers report that the kinds of tests that are used in

high-stakes contexts are less useful than other measures for helping them adjust their instruction. Many teachers also believe that most large-scale tests do not adequately measure students' skills and knowledge, and they are particularly concerned about the tests' validity for English language learners and students with special needs.

As high-stakes testing become more prevalent, researchers are beginning to collect more comprehensive information on teachers' and other educators' attitudes about high-stakes testing. This information is valuable to those responsible for developing and implementing test-based accountability systems, particularly because educators' attitudes and support for the testing system are likely to affect their decisions about whether or how to change their practices in response to those systems.

Effects on Student Achievement

Perhaps the most important policy question surrounding the use of high-stakes tests is the question of whether they lead to improved student learning. Regardless of how teachers or other educators respond, the primary rationale for attaching serious consequences to test scores is that it will lead to higher achievement. The adoption of new accountability requirements is often accompanied by claims that states that had strong test-based accountability systems in place experienced large and rapid achievement gains. However, determining whether such policies did, in fact, improve achievement is complicated by the fact that they often result in test-score inflation that distorts the meaning of scores on the high-stakes measure. In addition, test-based accountability policies are typically implemented in the context of other reforms, and it is not always possible to disentangle the effects of accountability from the effects of other initiatives. Nonetheless, several studies have tried to assess the impact of high-stakes testing on student achievement using tests other than the high-stakes tests. The results are decidedly mixed. Recent studies using National Assessment of Educational Progress state-level scores as the outcome suggested some positive relationships between high-stakes testing and student achievement, but these studies raise questions about the sources of those relationships, and they do not support strong causal conclusions.

One finding that is quite consistent, however, is that gains on state tests far outpace gains on low-stakes

measures of the constructs measured by the high-stakes test. As noted earlier, one of the challenges in evaluating the effects of high-stakes testing is the likelihood that the scores on the tests used in high-stakes testing systems will become distorted as a result of the high stakes. This phenomenon is known as *score inflation*, a term that refers to gains in test scores that are not accompanied by commensurate gains in the construct the test is designed to measure. The discrepant trends that are typically observed when comparing high-stakes testing gains with gains on low-stakes measures provide evidence of score inflation. If gains on a high-stakes mathematics test, for example, actually reflected an increase in students' mathematical knowledge and skills, their scores would be expected to rise on another mathematics test, assuming a reasonable degree of overlap in content.

The detection of score inflation is not as simple as comparing two trend lines, however. Gains on a high-stakes test may fail to generalize to another measure of the same construct for a variety of reasons, including substantive differences in the definition of the construct. For example, the fact that the National Assessment of Educational Progress score gains do not correspond perfectly to state high-stakes test-score gains undoubtedly reflects, to some degree, the different sets of standards and specifications used to construct each test. Moreover, a conclusion that scores on a particular test are inflated depends, in part, on one's definition of the construct being measured—for example, whether mathematics achievement is thought to encompass only the material included in a state's published standards or whether it is considered to be broader. Still, the extremely large discrepancies that are typically observed strongly suggest the presence of score inflation regardless of how narrowly one chooses to define the construct of interest.

There are also questions about how high-stakes testing affects members of specific groups, such as low-income students or English language learners. Teacher and principal surveys suggest that NCLB's requirement for reporting separately by subgroups has led to an increased focus on the achievement of students who in the past had sometimes been ignored. But there is also evidence that district and school personnel may try to keep the scores of low-performing students from counting, by taking steps such as retaining students in grade or classifying students as having certain disabilities that exclude them from testing. Such steps are more difficult to take in the context

of many NCLB-inspired accountability systems that require almost every student to be included, but they represent a possible risk stemming from high-stakes testing policies. Researchers have also expressed concerns that test-based graduation or promotion policies will lead to increased dropout rates, but the evidence on that outcome is inconclusive.

Ensuring Appropriate Use of High-Stakes Testing

High-stakes testing has increasingly become a central feature of the public education system in the United States, and public school students can now count on encountering dozens of high-stakes tests throughout their years in school. Similarly, nearly every teacher, principal, and district superintendent working in a public school or district is affected by these tests in some way. Because of their widespread use and impacts, it is critical that steps be taken to promote appropriate use of these tests and the information they produce. The National Research Council Committee on Appropriate Test Use pointed out that the value of high-stakes testing must be weighed against any potential negative consequences arising from such test use and against other information that could be used to make similar decisions. It is possible, for instance, that the use of a test to make a decision about course assignment will lead to certain negative consequences but that the consequences that would result from making that decision without the test-score information would be worse. Each high-stakes testing system should be evaluated in light of the particular context in which it is situated, and the validity and appropriateness of each specific use of test scores arising from that system should be carefully considered.

Several organizations have published guidelines or standards for appropriate use of tests under high-stakes conditions. Perhaps the best known of these are the *Standards for Educational and Psychological Testing*, published in 1999 by the American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education. One of the most important and widely discussed standards is that a single test score should not be used for important decisions but rather should be combined with other information. This standard, which is often described as a requirement for multiple measures, has been interpreted in different

ways. The spirit of the standards requires a compensatory system, in which low scores on one measure can be offset by good performance on other criteria, but that has not been followed consistently. Providing examinees with multiple opportunities to take the test (or equivalent forms of it) is consistent with the standards but is not generally considered sufficient to meet the requirement for multiple measures.

The standards also require evidence of the test's validity for the specific purpose for which it is being used. A test that has been shown to promote reasonably valid inferences when used for one purpose might not necessarily be appropriate for other purposes. There is often a temptation to use a single test for multiple purposes, particularly when educators believe they are forced to devote too much time to testing. As a result, schools and districts often use tests to make decisions that those tests were not necessarily designed to support. State accountability tests, for example, are now frequently used to make decisions about course placement even though states and districts have not gathered the necessary evidence to ensure such decisions are sound and lead to desirable outcomes.

Validity evidence should demonstrate that a particular test measures what it is intended to measure and that scores are not unduly influenced by extraneous (or "construct-irrelevant") conditions or attributes of the examinees. The likelihood that scores will be strongly affected by construct-irrelevant factors is of particular concern for students with disabilities and for students who are taking a test in a language other than the language they speak at home. As a result of NCLB's requirements for nearly universal student participation in testing, students in both of these groups are increasingly being included in state testing programs. States and other developers and users of tests bear the responsibility for ensuring that the tests are appropriate measures of the intended constructs for these students and that any accommodations or modifications made to the tests do not adversely affect the tests' validity.

One factor that influences the validity of information from high-stakes testing programs is the way scores are reported. This is particularly true when stakes are attached to the attainment of a specific target. In such cases, the method used for designing a performance measure can create incentives that encourage certain behaviors and discourage others. The NCLB approach to measuring AYP provides a good example. It provides incentives for educators

to pay attention to the achievement of traditionally low-performing subgroups of students, such as those with disabilities. Because it relies on the percentage of students scoring above the "proficient" threshold, it also creates incentives to focus on students who have the highest probability of moving above that threshold. Also, because it relies, for the most part, on scores at a single point in time rather than measuring gains, some educators have argued that the targets are unrealistic, an attitude that, if widely held, could adversely affect the quality of teachers' responses to the accountability provisions.

Another important consideration when evaluating the appropriateness of high-stakes testing is whether students have had sufficient opportunity to learn the material on which they are being tested. This is particularly important for tests with serious consequences for students, such as those that determine whether a student receives a high school diploma. For accountability tests that have consequences for educators rather than students, the issue of sufficient opportunity to learn is not as clear-cut, but there is still a need for policymakers who design those testing systems to ensure that educators have access to the resources (e.g., curriculum materials) needed to meet the goals that are set up for them.

Perhaps the most critical point to keep in mind when designing a high-stakes testing system or evaluating the effects of such a system is the well-known maxim that "what you test is what you get." When incentives are attached to test scores, it is likely that the behaviors of test-takers and those who educate them will be shaped by the specific content and format of the test, often in ways that test developers and users might not anticipate. Efforts to monitor the effects of high-stakes testing systems can help mitigate negative outcomes and can inform future test and accountability system development.

Laura S. Hamilton

See also Assessment; Criterion-Referenced Testing; No Child Left Behind; Standardized Tests; Validity

Further Readings

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing.* Washington, DC: Author.

Baker, E. L., Linn, R. L., Herman, J. L., Koretz, D. (2002*). Standards for educational accountability systems* (Policy Brief No. 5). Los Angeles: Center for Research on Evaluation, Standards, and Student Testing.

Center on Education Policy. (2006). *From the Capitol to the classroom: Year 4 of the No Child Left Behind Act.* Washington, DC: Author.

Hamilton, L. S. (2003). Assessment as a policy tool. *Review of Research in Education*, 27, 25–68.

Herman, J. (2005). *Making accountability work to improve student learning* (CSE Tech. Rep. No. 649). Los Angeles, CA: Center for Research on Evaluation, Standards, and Student Testing.

Heubert, J. P., & Hauser, R. M. (Eds.). (1999). *High stakes: Testing for tracking, promotion, and graduation.* Washington, DC: National Academy Press.

Koretz, D. (2003, April). *Attempting to discern the effects of the NCLB accountability provisions on learning.* Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Koretz, D., & Hamilton, L. S. (2006). Testing for accountability in K–12. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 531–578). Westport, CT: American Council on Education & Praeger.

Linn, R. L. (2000). Assessments and accountability. *Educational Researcher*, 29(2), 4–16.

Stecher, B. M. (2002). Consequences of large-scale, high-stakes testing on school and classroom practices. In L. S. Hamilton, B. M. Stecher, & S. P. Klein (Eds.), *Making sense of test-based accountability in education* (pp. 79–100). Santa Monica, CA: RAND.

HISPANIC AMERICANS

Hispanics, also known as Chicanas/os or Latinas/os, are the largest ethnic minority group in the United States. In 2001, the U.S. Census reported that Hispanics constituted 12% of the U.S. population, with 66% of these identifying as Mexican-heritage. However, this percentage underestimates the large number of Hispanics that enter the United States illegally and are routinely missed by census workers. By 2015, Hispanics will constitute the majority school-age population in the southwestern states, and dramatic increases in the school-age population of Hispanics in the South and the Midwest will be evidenced. Hispanic students are receiving much attention by researchers, educators, policymakers, and the media because, with the exception of the first wave of Cuban immigrants who came to the United States after Fidel Castro rose to power in 1959, as a group,

Hispanics have the highest rates of academic difficulties and school dropout of any ethnic group in the United States. They are also severely underrepresented in higher education and in professional occupations.

Despite controversy, many states concerned with Hispanic students' educational performance have increased funding and outreach to enroll young children in preschool enrichment programs such as Head Start and offer Hispanic families English classes and tips and strategies for supporting their children's schooling. Yet, many Hispanic children not only enter kindergarten with lower literacy and numeracy skills than their European- and Asian-heritage classmates, but they are also more likely to be held back from promotion to first grade. As they move through the K–12 grades, the gap between the educational accomplishments of these three ethnic groups continues to widen. Although African-heritage students also have poorer academic trajectories than European- and Asian-heritage students, as a group they are doing better academically than Hispanic-heritage students. Only Native American-heritage students perform as poorly or more poorly than Hispanic students in the K–12 grades; Native American-heritage students are also even less likely to attend and graduate from college and pursue graduate education and professional degrees.

Hispanics are also underrepresented in higher education and are less likely to have college-based educational goals than are other immigrant groups. For example, in California, the state with the largest Hispanic population, Hispanics constitute 40% of high school graduates, but only 6% are academically eligible to attend the University of California, the most prestigious public university in the state. Moreover, although recent statistics suggest that the number of Hispanics enrolled in college has increased dramatically in the past few decades, more careful analyses of these data have shown that most of this growth has occurred in the community college population and that a large proportion of Hispanic community college students either do not complete their degrees or, when they do, often do not transfer to 4-year universities.

This bleak educational picture is troubling to southwestern states with high proportions of Hispanic children and youth because it forecasts a future in which these states will lack a sufficiently trained workforce to meet their needs for skilled labor and professionals. As businesses that traditionally employed unskilled and semi-skilled workers, such as canneries, factories, and electronic, garment, and other types of assembly plants, move to Latin American and Asian countries that provide cheaper labor and allow them to compete with Asian imports, states that have a high density of Hispanics also worry about how their already strained welfare systems will be able to absorb the steady stream of Hispanic immigrants, many of which enter the United States illegally and are the subject of charged debates at the local, state, and federal levels.

Research and policies aimed at understanding and improving this bleak picture of Hispanic students' educational performance and futures must first acknowledge the diversity of the Hispanic population in the United States and the diversity of Hispanic students' educational trajectories. The low grades and achievement test group averages mask the success of a large number of Hispanic-heritage students who excel in school, attend prestigious universities, and become prominent professionals and politicians. Indeed, a major research trend since the 1990s has been to study the factors that promote the academic success and professional trajectories of Hispanic students. The following sections present an overview of the sources of diversity in Hispanic students' academic achievement and educational pathways and summarize relevant theory, research, and practice.

Who Are the Hispanic Students?

The label *Hispanic* includes individuals from Central and South America and the Caribbean Islands. According to the 2001 U.S. Census, the largest groups of Hispanics identify as Mexican heritage (66%, 22 million), Cuban heritage (4%, 1.3 million) or Puerto Rican heritage (9%, 3 million). There are an additional 3.8 million Puerto Ricans living on the island that became a U.S. territory in 1989 and who are thus also U.S. citizens. Social class is a central source of diversity in Hispanic students' academic performance and educational aspirations, because it plays an important role in the educational values and practices these students bring to school. Considerable research has shown that students whose family values and goals and social interaction and other behavioral practices match those of the school perform better than do students for whom these dimensions of home and school are discontinuous or in conflict. Because U.S. schools typically emphasize middle- and upperclass values and behavioral practices, poor and working-class students are often at a disadvantage and must learn new expectations and behaviors when they

come to school that their middle- and upper-class peers already take for granted.

The poor academic performance of Hispanic students relative to other ethnic minority groups is due, at least in part, to the fact that they are more likely to come from poor and working-class recent immigrant families in which parents have low levels of education and who, despite their high educational aspirations for their children, lack information about how to help their children succeed in U.S. schools; for example, information about the requirements for high school graduation and attending college. The burden of communicating this information is often placed on teachers and counselors in underresourced schools or on the students themselves, many of whom lack the cognitive sophistication necessary to chart their educational future. Cuts in federal and state funding for academic enrichment or outreach programs for students from ethnic groups with low high school graduation rates that are also underrepresented in higher education provide further challenges to the educational futures of Hispanic students.

The problematic educational future of Hispanic students and the political and public outcry and ambivalence toward helping these students succeed in school are due, at least in part, to the continued influx of Hispanic immigrants into the United States that has increased dramatically in the past decades. Most of these recent immigrants are unskilled or semiskilled laborers and refugees leaving war and poverty in their home countries, most notably El Salvador, Nicaragua, Guatemala, Haiti, and the Dominican Republic. A much smaller group of immigrants includes white-collar and professional workers from South America, most notably Argentina, Chile, Uruguay, and Peru. These recent immigrants constitute approximately 30% of the current Hispanic population of the United States. Despite the differences in their country of origin, social class, and educational levels, these recent immigrants endure the sacrifice of leaving behind family, the dangers of immigration, and the hardships of trying to make it in a new country because they share the dream that America is a land of opportunity and that with hard work and education, one can achieve whatever one desires. Although the realities of living in poverty and in a country that is increasingly hostile to immigrants can tarnish this dream, many recent Hispanic immigrants still hang on to their view of education as the key route out of poverty. Their children also enter school sharing their

parents' high educational aspirations, but by the end of middle school many, especially boys, no longer expect to finish high school or attend college.

Mexican-heritage students have the highest rates of academic difficulties and school dropout and the lowest rates of completion of 4-year colleges of any Hispanic group in the United States. Yet, Mexican-heritage students are themselves diverse. Many U.S.-born Mexican-heritage students self-identify as Chicana/o to signal their allegiance to the political movement of resistance against mainstream American values that occurred between the 1960s and 1990s, when many Mexican-heritage individuals, typically youth and young adults in California and Texas, embraced their Aztec roots and repudiated the American values imposed on them by what they viewed as an illegal conquest by the United States of Mexican territories. Since then, Chicanas/os have been vocal about the lack of representation of Mexican-heritage history and culture in the U.S. K–16 curriculum. Their actions have led researchers, policymakers, and educators to recognize and investigate the importance of identity, motivation, and school climate in Hispanic students' school achievement and educational pathways. Research has shown that identification with their culture of origin and viewing ethnic identity as central to their sense of self can act as protective factors for Hispanic students and provide the resilience that they need to cope with racism and discrimination at school and the larger society. This protective function has emerged more consistently for girls than for boys, suggesting the need to consider gender-related processes and mechanisms in Hispanic students' resilience and academic performance and pathways.

Importance of Geographical Location

Hispanic students' geographical location in the United States must also be considered when describing and explaining their school performance and pathways. Most Cuban- and Caribbean-heritage students live in Florida, most Puerto Rican–heritage students live in the Northeast (most notably New York City), and most students of Mexican American and Central American heritage live in the Southwest. Evaluations of Hispanic students' academic performance and trajectories must thus clarify the geographical region of focus as well as disaggregate the performance of Hispanic students in different urban and rural school settings. The neighborhood contexts of large urban

school districts, such as New York City and Los Angeles, also need to be considered because, as argued so convincingly by Rubén Rumbaut, Alejandro Portes, and others and captured by their concept of *segmented assimilation*, these receiving neighborhood contexts determine the slice of American values immigrants are exposed to and come to regard as mainstream. Hispanics attending public schools in Hollywood and wealthy suburbs of New York City will experience different school climates, American values, and statuses within their peer group than Hispanics attending school in East Los Angeles and Spanish Harlem. Hispanic students who attend public school in wealthier neighborhoods will more than likely be a minority, be less likely to be exposed to historical and cultural aspects of their heritage, may consider ethnicity less central than other dimensions of identity, and may be less aware of issues of privilege and discrimination than their Hispanic counterparts who attend school in less privileged, resourced schools. Yet, within these schools, in which Hispanic students are often the majority, they may achieve the power and status within their peer group that they lack in more privileged settings, and this positive ethnic identity and status may serve as a protective factor that promotes resilience and academic achievement for those students who manage, with the help of their family, friends, and teachers, and community agents, to sustain their dreams of attending college and attaining professional careers.

In the Midwest, Chicago has historically had the largest Hispanic population, but several midwestern states are experiencing dramatic increases in their Hispanic-heritage school-age population as a result of Hispanic families seeking jobs in the meat packing industry. The South, especially Georgia, North Carolina, and South Carolina, are also experiencing dramatic increases in the Hispanic school-age population as Hispanic families migrate to these states from initial entry points of Florida, Texas, and California in search of employment and less-hostile attitudes than those they encountered in these three receiving states. These three states continue to be the typical entry points not only because of their proximity to the Caribbean and Mexico but also because like other immigrant groups, Hispanics immigrants often join family and friends who are already in residence. These social networks are crucial for obtaining employment, housing, and informing parents about the requirements of school registration, medical services, and other resources that

are essential for children and adolescents' success in school. We now consider how immigration and variations in attitudes toward school associated with their country of origin affect Hispanic students' patterns of educational performance and achievement.

Diversity Due to Hispanics' U.S. Immigration History

Most recent Hispanic immigrants came to the United States either voluntarily in search of a better economic and educational future or as political refugees from the continued political strife in Central America, South America, and the Caribbean. In contrast, many Mexican-heritage students in the Southwest are descendants of the Mexican families who became U.S. citizens in 1848 following the Guadalupe Hidalgo treaty that annexed the southwestern states to the United States and displaced them from their lands and made them an often unwilling part of the colonial labor force. Descendants of these families are considered involuntary immigrants because they did not choose to become U.S. citizens; identification with the Chicana/o or the Latina/o label is one indication of this resistance and involuntary status. Puerto Ricans who have been active in the movement to grant Puerto Rico sovereignty from the United States also consider themselves involuntary immigrants. As proposed by John Ogbu and his followers, their perception of voluntary and involuntary status can affect Hispanic students' academic achievement and aspirations.

In their landmark ethnographic and case studies, John Ogbu, Marcelo Suarez-Orozco, and their followers found that often, voluntary immigrants perform better in school and have higher educational aspirations than involuntary immigrants. The better educational performance of recent voluntary immigrants presents an educational paradox because these foreign-born Hispanic students have fewer educational resources than their long-term U.S.-born Hispanic classmates. For example, especially in the case of Caribbean, Mexican, and Central American families, parents often have limited education, typically some elementary school or at most middle school; do not speak or read English; live in dangerous, low-income neighborhoods; and work long hours in menial jobs that prevent them from supervising their children, helping with homework, and participating in parent-involvement activities at school. Researchers that built on these landmark ethnographic

and case studies have shown that these voluntary immigrant students' strong cultural identification with their home countries, the centrality of their ethnic identity, and their determination to succeed academically to repay their parents' sacrifices are important assets that help them succeed academically. Their older siblings are also important "culture brokers" that socialize them into their new school system, help them with homework, and often sacrifice their own educational goals to help their younger siblings succeed academically.

Although it has confirmed the educational assets of recent voluntary Hispanic immigrants identified by qualitative researchers, quantitative research that has included larger, more representative national samples, such as the National Education Longitudinal Study of 1988, has also revealed the diversity of voluntary immigrant Hispanic students' academic performance. Many of these recent immigrants experience academic difficulties and school dropout and are performing more poorly in school than their U.S.-born Hispanic classmates.

Immigration History and Generation of Immigration

Regardless of whether Hispanic students are voluntary or involuntary immigrants, the historical period in which their families emigrated to the United States and also their generational status must be considered when assessing Hispanic students' school performance and educational goals and pathways. For example, the first wave of Cuban immigrants typically were middle- and upper-class well-educated professionals, whereas the Cuban families who emigrated to the United States in the 1980s as part of the *Mariel Boat Lift Amnesty*, known as *Marielitos*, and those who emigrated in the early 1990s due to economic hardship brought about by the disintegrating Soviet Union's reduction of its support of the Cuban economy were mostly low-income and poorly educated. Although they typically had to leave their material possessions and wealth in Cuba and, as is commonplace for immigrants, their professional degrees were not acknowledged in the United States, their middle- and upper-class origins and educational credentials were important assets for this first wave of immigration Cubans. First-wave-of-immigration Cubans currently are well-respected middle- and upper-class professionals and politicians, and their children have

performed so well academically that they are no longer considered by the U.S. federal government to be underrepresented in higher education. In contrast, *Marielitos*' educational trajectories are more problematic because they typically lack the educational and social-class know-how or *capital* of first-wave Cuban immigrants.

Like the *Marielitos* and subsequent Cuban immigrants, the majority of recent immigrants from Central America, Mexico, and other Caribbean nations typically are low-income and poorly educated families who came to the United States in search of work and a better education for their children. They often settle in rural areas in the Southwest or in dangerous, low-income neighborhoods in large cities, such as East Los Angeles and Spanish Harlem in New York City. The majority of Mexican immigrants are from the states of Jalisco, Michoacán, and Guanajuato, and Central American immigrants typically come from Guatemala, El Salvador, Nicaragua, and, increasingly, Honduras. Unlike the Caribbean refugees, many of these recent immigrants are undocumented, and their undocumented status affects their schooling by creating a climate of fear and danger that often leads to frequent moves and changes in schools, well-documented correlates of poor academic performance and school dropout. Also, although their undocumented status does not prevent these students from attending public K–12 schools, it limits their chances for enrollment in community colleges and public universities because many require social security numbers for admission and because only U.S. residents and citizens are eligible for state and federal financial aid packages. Although some states are currently considering legislation that will allow long-term undocumented students who complete high school in the United States to enroll in community colleges and 4-year state universities, this proposed legislation is not without controversy.

Despite their low economic and educational capital, their parents' high educational aspirations and support and the students' sense of obligation to their families are important educational assets for recent Hispanic immigrant students. Their high educational goals and aspirations lead many of these families to prioritize education over chores and other family-sustaining activities and motivate their children to persist in school despite poor grades or negative experiences such as racism and discrimination or low academic achievement. These sources of external

motivation and, in particular, their sense of obligation to their families may play a more important role in Hispanic students' persistence in school than the intrinsic sources of motivation often associated with European- and Asian-heritage students' educational achievement. Researchers have shown that high-achieving Hispanic students who enroll in prestigious private and public universities often attribute their success to their families, want to graduate to make their families proud, often send part of their financial aid monies home to their families, and are committed to serving as educational role models for their younger brothers, sisters, and other relatives. Of the ethnic minority students enrolled in college, Hispanics are the most likely to attribute their academic success to the mentoring of an older sibling. Many of these students, especially females, also state that whether or not they had attended college, their older siblings convinced their parents to allow them to attend college even when the college of their choice required that they live away from home.

Not all immigrant Hispanic families prioritize education over activities that sustain the family. As the realities of living in poverty in a country that is increasingly hostile to immigrants take root, parents may begin to rely on older children to help with chores and obtain employment to help the family's survival, thus negatively affecting their children's educational performance. As their children proceed through the elementary and middle school and perform poorly academically, many immigrant Hispanic parents also abandon their high academic aspirations for their children in favor of expectations that they view as more realistic because they do not require college degrees and are attainable during or soon after high school graduation, such as beautician, mechanic, or sales clerk. Parents view these occupations as acceptable because they are often better than the menial jobs they themselves perform and are steady sources of income that will allow their children to break the cycle of poverty and support their own families.

Second- and third-generation Hispanic students, especially males, often have a more skeptical view of education as a way out of poverty. As they move through adolescence and acquire the cognitive sophistication to understand issues of hierarchy and privilege, these students become aware that because of racism and discrimination, their educational credentials will not accrue them the same opportunities accrued by their European-heritage and, to an extent, Asian-heritage peers. Research has shown that for some Hispanic students, this awareness leads them to disengage from school and construct an identity that views succeeding in mainstream racist institutions, such as school, as undesirable. In contrast, other Hispanic students use their awareness and experiences of racism and discrimination as motivators to succeed in school and prove the racists and gatekeepers wrong. More research is needed to identify the characteristics that differentiate adolescents who pursue positive or negative educational pathways as a reaction to awareness of privilege and discrimination in society.

Hispanics in Higher Education

Although Hispanics are severely underrepresented in higher education, their participation in community colleges and universities has improved. Increases have been sharper for first-generation students; that is, students who are the first in their family to go to college. These students often commute to local community colleges or 4-year state universities and often work part time or full time to pay for their studies or augment loans and financial aid packages. As a group, these first-generation students are more likely to drop out of college after their first year or to take longer to finish their degrees than non-first-generation students. Reasons for their more challenging educational trajectories include their often weaker high school academic preparation, their work and family obligations that prevent them from taking advantage of educational supports offered by their college, and their higher propensity for feeling that they do not belong at the university, a feeling brought about by their difficulties establishing a supportive group of college peers. Establishing a supportive peer and friend group at college takes time and energy, and the academic, work, and family demands experienced by many first-generation students make it difficult for these students to develop supportive college friendships. Yet, considerable research has shown that these supportive peer networks are crucial for the adjustment and academic success of first-generation college students.

Hispanic students who are not the first in their families to attend college are more likely to do well and graduate. Many of these students are the children of middle- and upper-class Hispanic immigrant or U.S.-born parents who are themselves college graduates. Their children enter college with adequate academic

high school preparation and unburdened by work and family obligations because their parents are able to finance their educational and living expenses. Thus, their academic trajectories resemble those of their European- and Asian-heritage peers. Because past research has emphasized the educational trajectories of Hispanic students who are the first in their families to attend college, not much is known about the college trajectories of Hispanic students from middle- and upper-class professional families. Because these students represent a substantial proportion of Hispanic college graduates and professionals, investigating the factors that promote their success will be important for improving the educational trajectories of poor and working-class Hispanic students.

Margarita Azmitia

See also Cultural Diversity; Ethnicity and Race; Family Influences; Home Environment and Academic Intrinsic Motivation; Immigration; Parental Expectations

Further Readings

Azmitia, M., & Cooper, C. R. (2001). Good or bad—Peer influences on Latino and European American adolescents' pathways through school. *Journal of Education for Students Placed at Risk, 1–2*, 45–72.

Contreras, F., & Gándara, P. (2006). Latinas/os in the Ph.D. pipeline: A case of historical and contemporary exclusion. In J. Castellanos & A. Gloria (Eds.), *Journey to a Ph.D.: The Latina/o experience in higher education.* Sterling, VA: Stylus.

Contreras, J. M., Kerns, K. A., & Neal-Barnett, A. M. (Eds.). (2002). *Latino children and families in the United States: Current research and future directions.* Westport, CT: Praeger.

Gibson, M., Gándara, P., & Koyama, J. P. (2004). *School connections: U.S. Mexican youth, peers, and school achievement.* New York: Teachers College Press.

Miller, L. S. (2005). Exploring high academic performance: The case of Latinos in higher education. *Journal of Higher Education, 4*, 252–271.

Pedraza, P., & Melissa, R. (Eds.). (2005). *Latino education: An agenda for community action.* Mahwah, NJ: Lawrence Erlbaum.

Reese, L., Kroesen, K., & Gallimore, R. (2000). Agency and school performance among urban Latino youth. In R. Taylor & M. Wang (Eds.), *Resilience across contexts: Family, work, culture, and community* (pp. 265–332). Mahwah, NJ: Lawrence Erlbaum.

Suarez-Orozco, M. M., & Páez, M. M. (Eds.). (2002). *Latinos: Remaking America.* Berkeley: University of California Press.

Valenzuela, A. (1999). *Subtractive schooling: U.S.-Mexican youth and the politics of caring.* Albany: SUNY Press.

Zentella, A. C. (1997). *Growing up bilingual: Puerto Rican children in New York.* New York: Blackwell.

HIV/AIDS

In 1981, medical literature first described an unusual constellation of symptoms in five young homosexual men in Los Angeles. Subsequent reports described similar symptoms in residents of other cities, such as San Francisco and New York, not only among men who were having sex with men (MSM) but also among individuals who injected drugs and in people with hemophilia. These cases suggested that the cause was in the blood and was transmitted by sexual contact, sharing infected drugs, or receiving blood donations. In September of 1982, the Centers for Disease Control and Prevention (CDC) published a case definition using the current name *acquired immune deficiency syndrome (AIDS)*.

It is now known that AIDS is caused by the human immunodeficiency virus (HIV). The virus can be transmitted when infected blood or bodily fluids (including semen and vaginal secretions) gain access to the bloodstream of an uninfected person. HIV can also be transmitted to babies during pregnancy or delivery and through breast-feeding.

Being HIV-positive is not the same as having AIDS. A person with HIV can look and feel healthy for many years. Over time, HIV weakens the immune system, and the ability of a person to fight ailments diminishes. Consequently, persons living with HIV/AIDS are susceptible to "opportunistic" infections such as bacteria, protozoa, fungi, and viruses. HIV becomes AIDS when the CD4+ cell count drops below 200. The progression of HIV to AIDS is different in every person. New medications have revolutionized treatment for HIV/AIDS and have extended the life of people with HIV/AIDS. These medications do not cure HIV or AIDS; rather, they delay the progression of the disease and treat opportunistic infections. AIDS became the leading cause of death for people between the ages of 25 to 44 years old, although new medications have cut the AIDS death rate dramatically.

In the United States, the CDC estimates there are 1 to 1.2 million people living with HIV/AIDS; about one quarter of these people do not know that they

have the virus. Men represent 75% of HIV infections. In the past, transmission was mainly via MSM and injection drug use, but in recent years heterosexual intercourse has become the predominant mode of transmission. Biologically, women are more susceptible than men to heterosexual infection, and HIV is increasing more rapidly among women than men. Another change is the racial and ethnic distribution of HIV. Initially HIV/AIDS disease affected mainly White Anglos; now the vast majority of new cases in the United States occur among Blacks, Hispanics, and, more recently, Asians. Few new cases are attributed to blood transfusions because of safety measures that have been implemented. A few studies show that there are different strains of HIV. Right now, there is much evidence of HIV1, and cases of HIV2 are rare. With no cure for HIV/AIDS in sight, scientists are working to develop a vaccine.

In the absence of a cure or vaccine, the best hope for combating the HIV/AIDS epidemic is prevention. It is recommended that people who are sexually active use protection for every sexual act, and get tested periodically. In totally monogamous couples who have both tested negative after an extended period of 100% precaution, it is recommended to use protection for every sexual act outside the couple. The use of clean needles is recommended for people using intravenous drugs. A variety of behavioral interventions have been developed for specific risk groups. Definitions of "risk" extend beyond the spheres of the individual and the couple and include contextual factors such as community and societal factors. Poverty, discrimination, substance use or abuse, and violence have been recognized as risk factors, particularly in communities of color.

Claudia L. Moreno

See also Abstinence Education; Drug Abuse

Further Readings

Blank, M. B., & Eisenberg, M. (2007). HIV and mental illness: Opportunities for prevention. *Journal of Prevention and Intervention in the Community*, *33*(1–2), 1–4.
Kaiser Family Foundation. (2007, June). *HIV/AIDS policy fact sheet: HIV testing in the United States.* Retrieved August 13, 2007, from http://www.kff.org/hivaids/upload/6094–06.pdf
U.S. Library of Medicine & National Institutes of Health. (n.d.). *AIDS.* Retrieved April 2, 2007, from Medline Plus Web site: http://www.nlm.nih.gov/medlineplus/aids.html

HOME EDUCATION

Home education (U.K. term), or *homeschooling* (U.S. term), describes an education based in the home rather than in an institution and one that is facilitated by parents rather than teachers. Whereas a school education takes place in a specific location, home-based education is likely to have a more transient focus as parents and children make educational visits, attend social gatherings, and join other group activities outside the home. Though school is the most widely preferred form of education for the majority of families and governments, home education in the United Kingdom, the United States, and many other nations is a fundamental right that parents are at liberty to implement.

The Law

The U.S. Constitution does not regulate homeschooling; this is left to the individual states. According to the Christian organization Home School Legal Defense Association, approximately half of U.S. states have little or no regulation; the other half have moderate to high legislation. Those states with little regulation require parents only to inform them of their decision to homeschool, whereas states with moderate to high legislation require the submission of test scores. The age at which compulsory education begins for children varies between 5 years of age and 8 years of age depending on the state.

In the United Kingdom the law permitting home education makes home education equal in legislation with school education. Section 9 of Education Act 1996 requires that the parents of a child of "school age shall cause him to receive efficient full-time education suitable: (a) to his age, ability and aptitude, and (b) to any special educational needs he may have, either by regular attendance at school or otherwise." In the United Kingdom there is no clear definition of "an education" and no clear guidance from the Department of Education regarding what an education should look like. This can lead to confusion and misunderstanding between home educators and their local authorities. More recently in the United Kingdom, the Children Act 2004, strengthened by Statutory Guidance 2007, has placed a duty on local authorities to safeguard children's welfare, and the distinction between welfare and education has become

blurred. The Department for Education and Skills (DfES) has issued home education guidelines for local authorities, but these are contentious within home education organizations and among some local authorities.

Numbers

Numbers of home educators are hard to ascertain. In the United States, whereas some states count home-schoolers closely, other states are less rigorous. The U.S. Department of Education estimates that in 2003, there were approximately 1.1 million homeschooled students, representing 2.2% of the school-age population. This demonstrates a growth from 1999, when the figures were 850,000 and 1.1%, respectively.

In the United Kingdom there is no requirement to register as a home educator with any formal administrative body, and therefore, numbers are impossible to calculate with accuracy. In the United Kingdom estimated numbers vary between 10,000 and 170,000. Mike Fortune-Wood, home education activist and researcher, used a freedom of information request to arrive at his conclusion that the figure of known home-educated children may be about 18,100. By extrapolating to include children not known to their local authorities, Fortune-Wood suggests the population may be close to 45,250 children, a figure that represents about 0.5% of the national school-age population. Recent 2007 figures broadly support these figures. This latter research suggests that numbers of home-educated children have increased threefold since 1999.

Rapid Growth

Whereas there is evidence of a rapid growth in home education over the past 10 years, there is less clarity about the reasons behind this. Some sources blame alleged poor standards in schools; others cite the popularity of individualism. In the United States, where there are many religious families choosing to home-school, the decision to homeschool may be strongly related to the family's religious connections. In the United Kingdom religiously orientated home educators are in the minority, and home education is more likely to be attributable to either bullying in school or simply a lifestyle choice. Despite coming to home education from diverse starting points, parents who persist often come to see home education as a positive

step and enjoy the flexibility that home education can bring.

From Birth or Withdrawal From School

Press coverage tends to focus on children withdrawn from schools; however, such children appear to account for only half of all home-educated children. The "from birthers," as they have been called, represent a less vocal group as it includes confident parents who are less likely to approach support organizations and unlikely to come to the attention of the media. From the remaining 50%, half of these children are withdrawn during their primary years. Research by Paula Rothermel at the University of Durham has found that these children tend to have parents who previously considered home education but who were persuaded by friends and family to send their children to school. Thus, when an incident arises at school, the parents withdraw their children. This incident alone, however, is unlikely to be the base cause for the withdrawal. Only parents who withdraw their secondary-age children, accounting for 25% of the overall figure, tend to be those who never before considered home education but who decide to home educate following a serious incident such as bullying in school. This latter group is more likely to include parents who are less confident, less affluent, and often less educated than "from birthers" and "primary-years-withdrawal" parents.

Children With Special Needs

Research suggests that parents of children with special educational needs may face more opposition from their local authorities than parents of ordinary children. This stems from concerns that parents may not be able to provide the appropriate environment for their child. Parents, on the other hand, often take the view that they are best placed to decide whether or not they can provide what their child needs. In the United Kingdom, difficulties with unwelcome local authority intervention have led some parents to avoid the statementing process, whereas others have nonetheless welcomed the funding a Statement of Special Educational Needs (SEN) can bring. Rothermel's 2002 research found that 20% of parents who withdraw children from mainstream schooling do so because of what they describe as "mismanagement of SEN at the school." Tiny Arora, educational psychologist with

responsibility for home education in Sheffield, United Kingdom, has also found that parents who home educate withdraw their SEN children because of the school's failure to meet their child's needs.

Shift Over Time

In all groups there tends to be a shift over time in the family's approach to education and lifestyle. Generally, whatever the starting point, families tend to move toward less acceptance of societal norms. It is not uncommon for families to graduate to questioning health issues and to become more politically skeptical as home education progresses. Research in the United Kingdom shows that where initially the father may have been the main breadwinner or where both parents may have worked full time, once they begin to home educate, both parents may work part time to ensure division of labor in the home and time spent with the children. Arora has also highlighted a shift in focus as parents of children with special educational needs grow in confidence in their own ability to home educate.

Children in and out of School

Preconceptions about children sitting isolated at home are outdated, as are ideas about families steadfastly adhering to home education come what may. In reality, it is not unusual to find that a home-educating family will include children both in and out of school. The parents may start home educating one secondary-age child as a result of an incident at school, while continuing to school their other children. Alternatively, a child who has never been to school may decide to try school, and this might, in turn, lead to another sibling asking to go to school. Families who home educate all their children from birth or school age through to the end of compulsory education are in the minority. Home education is characterized by a motivated, reactionary, and child-led, but not arbitrary, style of parenting.

Academic Results

Both in the United States and the United Kingdom, research has found that home-educated children do well on standardized tests. Rothermel has found, moreover, that outside school, working-class children fare no worse academically than their more affluent peers. Of particular interest is the finding that some of the children achieving high scores learn in unstructured ways and with little, if any, work undertaken while sitting at a table. A DfES meta-study by Charles Desforges and Alberto Abouchaar in 2003 found that parental involvement is central to achievement after all other factors are taken into account. Therefore, it is hardly surprising that home-educated children, with their inherently involved, committed parents, do well. Just as home education correlates positively with academic outcome, it would be prudent to conclude that all children, in or out of school, with hands-on, interested parents are likely to do well. The finding that state-school children nationally do not do as well as their home-educated counterparts is likely to be more a comment on the standard of school education rather than a tribute to home education.

U.K. research has found that in school, cognitive challenges take up just 1% of lesson time. The U.K. Office of National Statistics has shown that the amount of time that parents spend caring for their child is an average of just 32 minutes a day. These data contrast with home education where cognitive challenges and parental input are characteristic features. Home education, with its individually paced, child-centered approach to learning, can provide time for children to absorb, assimilate, and understand challenging problems. School-based research suggests that when children are given the time to explore ideas for themselves and digest information at their own pace, they benefit more than they do in situations where the information is delivered didactically and soon forgotten. Karmilloff Smith in the United Kingdom has generated ideas about "incubation" periods, whereby children take time to absorb and unravel problems presented to them.

Social Skills

Research on both sides of the Atlantic Ocean shows that homeschooled students demonstrate good social skills. Rothermel's U.K. research has suggested that home-educated children are less inclined to manage criticism well, and this has been associated with the environment of positive attribution in which these children learn. In terms of socialization, as home education increases in popularity, so too do opportunities to interact with other home educators. In the United States homeschooling is a well-known phenomenon, and social contact between homeschooling families is

common. Similarly, in the United Kingdom most communities include home educators, although attendance at regular events is likely to involve traveling some distance. U.K. research shows that parents consider that their input in terms of ensuring opportunities for socialization is an essential factor in successful home education.

With the growth of activities before and after school, many home-educated children can join music groups, sporting events, Brownies, Guides, Cubs, Scouts, cadets, and community-based after-school groups, together with swimming, dancing, and drama groups. Although there will be some children who do not have access to such opportunities, home education is generally characterized by a level of parental commitment that makes inclusion in community groups with both home- and school-educated children increasingly likely. In the United States and the United Kingdom in particular, Internet networks bring home educators together for social gatherings and support. Together with the more traditional phone "tree" to disseminate notices and the growth in home education, home-educated children rarely need to be socially isolated.

The Process of Home Education

There are three basic styles of home education:

1. *Child-led.* These families tend to take each day as it comes and to follow the children's lead. Autonomous education, or "unschooling" as it is called in the United States, has the central tenet of child-directed learning (as opposed to child-centered), whereby children determine what, when, how, and why they learn. The children follow their own interests entirely, with encouragement and facilitation by parents.

2. *Mixed.* These families will adopt more of a shared approach to learning. The parents may expect some formal input from the children but will also follow the children's lead in giving them plenty of freedom to decide for themselves what they want to do.

3. *Parent-led.* It is more likely that these children will follow a purchased curriculum, and their parents will make all decisions about how the children's days are organized.

The overall format for home education is characterized by liquidity and flexibility. All three approaches tend to be child centered in that the education revolves around the individual child or children's needs. This child-centered ethos is unrelated to the degree of formality adopted. Families often use different approaches for different children; thus, whereas one child may prefer a workbook or curriculum-led education, another child may respond better to a more liberal style. Parents tend to adapt their approach depending on the child's age. Rothermel has found that families who have recently withdrawn children from school tend to follow a curriculum initially, relaxing this over time as they move to a child-led approach. In contrast, parents of children who have been home educated from birth tend to move from the informal to the formal, that is, the reverse of parents who have withdrawn their children from school. In approaching exams, children typically prefer a more formal path. About one fifth of British home educators maintain a clear structure to their children's home education, one half describe themselves as informal in their approach, and the remainder employ a mixed approach.

Families who home educate their children from age 5 to 16 years make up the minority and are often those with larger families. Within home-educating families, children who, over the compulsory education years, experience a school and home education mix, make up the majority (this would need retrospective research to fully establish this). North American data from Brian Ray suggest that homeschooled children are more interested in reading, politics, and community activities than are non-homeschooled children.

Deschooling

Deschooling is the term often used for the process of effectively "deprogramming" children following experiences at school. This term exists only among families who have removed a child from school, usually because of trauma of some kind. At one end of the spectrum, deschooling may take the form of children simply learning that they do not have to ask permission to go to the toilet, but more usually the term describes a period of up to a year during which children are given the space to "move on," following a period of unhappiness, such as having been bullied at school. Deschooling is more commonly associated with secondary-age children who find the shift from institutional, didactic education to a self-motivated home education more difficult than younger children do. The deschooling process is a therapeutic and philosophical approach that has little empirical research

attached to it although it is often discussed. For a fuller understanding of the underlying deschooling philosophy, the work most quoted is Ivan Illich's book, *Deschooling Society*.

Home Educators

The stereotypes of home educators are that either they are hippy rebels or hot housing fanatics. Typically, they are referred to as if they were a "type" of parent. In fact, home educators tend to have little if anything in common with each other beyond their decision to home educate and their desire to take full responsibility for their children. This desire to take control or responsibility is similar whether parents are Catholics, atheists, or Jews. However, the homogeneity stops there. Home educators experience between-family differences that will often prevent anything more than superficial friendships forming for the purpose of providing shared opportunities for their children (e.g., arranging educational visits and taking advantage of cheaper school rates). There are intergroup differences—for example, Muslim groups may join up with secular groups to co-provide educational activities for their children, but the two groups may have nothing else in common. The home education movement is made up of an eclectic mix of people whose overriding sense of responsibility leads them to relationships with those with whom they ordinarily would not share anything in common. This willingness to join forces to take advantage of educational rates on activities is a feature peculiar to home education. Home education is not a homogenous activity but is undertaken in different ways by different people, even within groups (e.g., religious groups). There are, however, national groups and national Web sites that add cohesion to the movement and give it a relatively powerful voice.

U.K. research shows that at least 20% of home-educating parents have no postschool qualifications. However, of those professional parents, the largest group (13%) of is made up of school teachers and college lecturers. Approximately 10% of home educators are employed in manual work, such as machinists, laborers, truck drivers, and factory workers. Research evidence suggests that less-educated parents who home educate their children often resume their own education with renewed vigor, once faced with the prospect of keeping pace with their children.

Home Education Around the World

Although home education exists in many countries, in Europe at least, the largest population of home educators is in the United Kingdom. Most European countries make home education possible, although in Germany it remains illegal. In Australia home education has seen rapid growth over the past decade. In New Zealand the education authorities have accepted that the education provided through home education is generally a good one, with over 90% of students reviewed considered to be taught at least as regularly and well as in a registered school. Moreover, in New Zealand parents can, if they wish, receive an annual stipend to home educate starting at about ₤290 for the first child and lesser amounts for subsequent children. The home education community in New Zealand represents a good mix of religious and secular home educators. However, the world leader is the United States, where the federal government has, in recent years, lent strong support to the homeschool movement.

Potential Problems

Pressures on Parents

British research shows that the most common pressures cited by families who home educate are other people's opinions, not being accepted in their community and the ensuing isolation that this brings, parents not having enough time for themselves, financial worries, the exhausting and time-consuming nature of home education, and the potential lack of resources. Some parents also say they feel pressured by the responsibility of home educating. Despite the pressures linked to home education, research suggests that home-educated children grow into capable adults and that many of them choose to home educate their own children.

Other People's Opinions and Sense of Isolation

The sense of isolation and lack of acceptance in the local community can cause children to suffer, making them feel awkward and shy. It is not uncommon to hear home-educated children complain about they way the feel they are being interrogated by strangers on the street during school hours. Some of this may be well-meaning interest, but too often it causes distress to children and increases their sense of being

different. Sometimes members of the public assume the children have special needs, and questions about reading and writing are common.

Lack of Time for Parents Alone

Parents of older children can sometimes negotiate with their children for some time alone, although this does remain a contentious issue where younger children are involved. Parents who continue to home educate come to accept lack of time for themselves as an inevitable aspect of home education. Many parents find it difficult to cope without sufficient personal time, and this can certainly place a strain on relationships within the family, who after all, are in much closer contact than would be the case in a family with children who go to school.

Financial Worries

When parents home educate their children, financial pressures tend to either remain, or be absorbed, into the culture of the family. Generally speaking, home-educating families live on one income and are less well off than families whose children attend school. However, 2003 DfES research shows that parental income is not a predictor of poor attainment. Rothermel's research has found that home-educating families tend to enjoy themselves. The value of happiness has been addressed by economists, such as Richard Layard, who have translated happiness into positive financial terms. Thus, although home-educating families are likely to be less well off financially, there may be some compensation in terms of the children's contentment and academic attainment. Moreover, U.S. research by Brian Ray suggests that homeschooled children are happier in their adult lives than schooled children.

Exhausting and Time-Consuming Nature of Home Education

Research from Australia by Rosanne Trevaskis in 2006 has identified maternal overload as a problem in home education because of the commitment and responsibility that so often fall to the mothers. However, in the United Kingdom both mothers and fathers express concerns over this issue. This pressure can be reduced when families get together and share some of the associated responsibilities. Some local groups

meet weekly, and parents can relax in a group care situation. Often home-educated children attend activities that allow the parents to take a break from parenting, although much depends on the ages, number, and needs of the children.

Lack of Resources

Concerns over resources are overcome as families share resources, ideas, discount cards, and so forth. The increasing use of e-mail discussion groups by home educators has led to a further sharing of resources.

Pressure to Perform

Home educators known to their local authorities often feel under pressure to demonstrate that their children are achieving to the standard that they would if they were at school. However, the list of research that has found home-educated children to perform as well as, and often better than, their schooled counterparts is growing. This body of research has consistently shown that home-educated children can do well both socially and academically. This is in spite of the often informal atmosphere in which such children learn. As yet, so far as I am aware, there has been no research that has shown home education to be detrimental to children's academic, social, and psychological well-being. Research with adults who were home educated as children demonstrates successful outcomes in terms of these adults' ability to make a positive contribution to society. Nevertheless, local authorities do sometimes put particular pressure on parents who feel less sure of their decision, and this pressure sometimes leads to parents ceasing to home educate. In the United Kingdom about 40% of families who home educate their children may not be known to their local authorities, and one of the reasons why families choose not to make themselves known is because of the pressure that so often comes from contact with local authorities.

Late Readers

Home-educated children may appear to make slower progress with their reading and writing than schoolchildren. This is usually because their parents prefer to let them learn at their own pace. This attitude is sometimes, but should not be, confused with schools' poor readers, whereby the children may be experiencing real problems with their reading. Many

home educators adopt this slower-pace approach to reading and writing, and there is no research evidence to suggest that this method, in the context of home education, leads to later problems. Home education–based research shows that late readers soon catch up with their earlier reading peers, often becoming keen readers. Nevertheless, for some parents, having a child aged between 6 and 10 years who cannot read, can and does lead to criticism of their decision to home educate and to labels of dyslexia. The research shows that these late readers become competent and avid readers.

Family and Friend Pressures

Rothermel's 2002 research shows that friends and families, particularly in the early days, can put home-educating parents under considerable pressure to meet standard expectations. As new home educators build friendships within the home-educating community, the force of this pressure diminishes, but initially it can cause a great deal of distress and some families decide not to continue with their home education. Disagreement between spouses over home education can be debilitating for the entire family. For example, the mother may want to home educate while the father needs further persuasion. Often there is an agreement to see how it goes for 1 or 2 years.

Children Labeled by Others as Hyperactive

Families who make an early decision to home educate focus their ideas in a different way than do families of children who are destined for school. For example, without the anticipation of starting school, there is usually no preparation for preschool, often little attempt to prepare for early writing and reading, and no pressing need to train the child to sit still and concentrate. Thus, at 4 or 5 years old, when school-children are usually able to write and read a little and sit for periods at their tables engaged in teacher-directed activities involving concentration, the home-educated child of this age can still be at a play-based stage, where he or she is not required to sit and learn or raise a hand to ask a question. It can seem that these children are hyperactive and unruly, unable to engage in proper activity. Not all home-educated children are like this, but for children from families who adopt the "learn at their own pace" philosophy, this is not unusual.

Formal Qualifications

In the United Kingdom, children who are home educated are able to take the same formal educational qualifications as schoolchildren. However, in practice this often requires strenuous efforts on the part of the parents in negotiating the minefield of taking exams outside school. In response to the growing demand for this service, numerous organizations have sprung up. The topic of access to qualifications has been the subject of a Westminster Adjournment Debate (May 13, 2003) at which the Minister said, "The playing field should not be uneven" for home-educated children wanting to take exams. However, at a second Westminster Adjournment Debate on home education (May 4, 2004), Sir George Young pointed out to the Minister that despite promises of a level playing field, the situation had not improved and that taking examinations for external candidates could be prohibitively expensive. Those who home educate their children during their secondary years report that there are plenty of resources available for children in this age group, although taking the General Certificates of Secondary Education (GCSEs) and A levels outside a college or sixth form can be expensive, and home-educated children often opt into the formal system part time at this point by entering a sixth form or further education college.

Divorced and Single Parents

There is evidence of families split by divorce still continuing to home educate. While this is easier and more stabilizing for the children when both parents continue to cooperate in home-educating their children, the available evidence indicates that the sole home-educating parent can, dependent on his or her support networks, continue to home educate with success. Although most home-educating families involve two parents, in the United Kingdom about 10% are single-parent families.

Transition to School

Transition to school is not known to cause problems. Research looking specifically at children's transition from home to school and from school to home education has found that children moving from home education into school tend to adjust well when teachers are professional and considerate. Parents who

decide to send their children to school usually believe that it is in the children's best interests. However, some parents choose to send their home-educated children to school through what they perceive as a necessity. This may be because one parent is to take on full responsibility for the children and needs to work, for example, following a separation. Home education can be suitable for some children, some of the time, and families are likely to follow their own and their children's needs, rather than steadfastly pursuing a home-educating agenda regardless of circumstance.

Criticisms

Critics of home education focus on the damage it will do to the children; for example, the lack of qualifications (in the British system), lack of exposure to diversity, absence of socialization skills, and the potential difficulty entering mainstream life. However, there is, as yet, no research evidence to support this viewpoint. In the United Kingdom concerns have been raised that home education could serve as a cover for child abuse, although the DfES has accepted in writing that it has no evidence to support this concern at this time. Nevertheless, concerns about abuse have been linked to the obligations set out in the Children Act 2004, and the DfES has come under pressure for lobby groups to tighten the regulations surrounding home education. There is evidence that some local authority officers and health workers are adopting the counterview that home education is itself a form of abuse. Nevertheless, it is a fact that in the United Kingdom and the United States, most children do attend school; not attending school is likely to attract fierce criticism from some quarters. Rothermel's research shows that the main disadvantage of home education is "other people's opinions."

In the United Kingdom home-educating families often report that they feel vulnerable to criticism by school-based educators on the issue of their children's socialization and assumed isolation. This exposure to criticism may well stem from the lack of information held by local authorities and government about home educators. The situation for home educators is that in the absence of any substantive evidence to the contrary, their situation remains open to speculation. As long as there remains a lack of standardization and cooperation in local authorities' approach to families, home educators are likely to resist registration and

investigation. At least in schools, evidence can be collated to support or dispute a claim, but this is not the case for home education.

Increasingly there are criticisms of home education as promoting separation anxiety. However, for a child to be diagnosed with separation anxiety he or she needs to meet the *Diagnostic and Statistical Manual of Mental Disorders* (Fourth Edition, Text Revision) *(DSM-IV-TR)* diagnostic criteria, which amount to more than a child's preference to be home educated. For diagnosis a child would be expected to exhibit "developmentally inappropriate and excessive anxiety concerning separation from home or from those to whom the individual is attached," and this would be evidenced by three different behaviors set out in *DSM-IV*.

Criticisms of home education research often center on the fact that much of the U.S. research uses Christian samples, and that in many cases, particularly in the United Kingdom, families are not randomly selected for participation because home educators are not required to register their home education nor undergo any summative assessment on a national scale; their data cannot, therefore, be plucked from a database. In contrast, school-educated children are both registered and tested, and their attainment data are stored electronically and can be accessed without the express permission of the children or their parents.

In the United States the prominent education commentator Michael Apple is a vocal critic of homeschooling, arguing that the commitment and motivation of homeschooling parents could be better used in making a positive investment in schools, rather than withdrawing their potential contribution to the common good as they do when they choose to homeschool. Similar theoretical arguments against homeschooling have been voiced by North Americans Chris Lubienski and Rob Reich.

Paula Rothermel

See also Alternative Academic Assessment; Early Child Care and Education; Learning; Learning Communities

Further Readings

Aurini, J., & Davies, S. (2005). Choice without markets: Homeschooling in the context of private education. *British Journal of Sociology of Education*, *26*(4), 461–474.

Hill, P. T. (2000). Home schooling and the future of public education. *Peabody Journal of Education*, *75*(1–2), 20–31.

Stevens, M. (2003). *Kingdom of children.* Princeton, NJ: Princeton University Press.

Home Environment and Academic Intrinsic Motivation

Academic intrinsic motivation (AIM) is defined as the enjoyment of school learning characterized by a mastery orientation, curiosity, persistence, task endogeny (i.e., pleasure in and orientation toward learning and task involvement), and the learning of challenging, difficult, and novel tasks. It concerns pleasure derived from the process of learning without receiving external or extrinsic consequences. Its significance for children's learning and development has been increasingly acknowledged over the years due to extensive research indicating that children with higher levels of AIM are more competent in the academic domain from early childhood through early adulthood. Such children show higher achievement on standardized tests as well as teacher-assigned grades, greater sense of academic competence, lower academic anxiety, less extrinsic orientation to learning, and higher educational achievement as young adults. AIM has also been shown to be related to aspects of home environment, indicating that the roles of parents and home stimulation are exceptionally important to facilitating its development. This entry elaborates on specific relations between home environment and AIM and discusses interventions for, and implications of, enhancing such academic motivation in children and adolescents.

To study the development of children's AIM, a psychometrically sound, published instrument was developed, called the Children's Academic Intrinsic Motivation Inventory (CAIMI), which divides AIM into four subject areas (reading, math, social studies, and science) as well as for school in general. Research employing this instrument, conducted within the Fullerton Longitudinal Study (FLS), is discussed as it has provided an empirical basis investigating the relationships between AIM and home environment. In addition, there has been other research on allied types of academic motivation indicating the importance of parents and the home environment; such findings are briefly described as well. This entry concludes with intervention suggestions based on the role of home environment in the development of children's AIM.

Conceptualizations

Academic Intrinsic Motivation

AIM, as defined in the previous section, is based upon three theoretical foundations, each having important implications for the role of environment in its development. These include cognitive discrepancy, competence/mastery, and attribution.

Regarding cognitive discrepancy theories, intrinsic motivation is viewed as the result of encountering stimuli that do not match existing cognitive structures, thereby creating motivation to reduce this discrepancy. Stimuli and learning materials that produce cognitive discrepancy can be expected to produce intrinsic motivation in the form of curiosity or exploration. Such stimuli would include those that are novel, complex, incongruous, surprising, and experiential.

Competence/mastery theories of intrinsic motivation concern children's experience of effectiveness in interaction with their environment. Children seek to interact effectively with their environment, and to the extent that they experience mastery, intrinsic motivation is enhanced. Central to this theory is the child's sense of being in control, that is, being a causal agent or influencing the environment by producing successful and noticeable outcomes. Activities may be considered broadly, including play, interactions with toys, learning materials, or individuals such as parents, peers, or teachers. Experiences that enhance competence/mastery intrinsic motivation include responsiveness of play materials and the social environment, parental provision of effectance feedback (i.e., competency information) to the child, and availability of materials and activities at an optimal level of challenge relative to the child's developmental abilities.

The attribution approach concerns the impact of extrinsic consequences for learning on intrinsic motivation. Provision of extrinsic consequences is external to the learning process and therefore not derived from learning per se. Such contingencies have important consequences for the development of intrinsic motivation because they affect the child's perception of the reasons for their engagement in an activity. If children perceive being engaged in an activity in order to receive an extrinsic consequence (e.g., money, toys), their focus of motivation is likely to shift from the

process of learning to the receipt of the outcome. Hence, their sense of intrinsic motivation would be reduced. However, if the extrinsic consequence augments the child's sense of competence, then intrinsic motivation is not likely to be adversely affected. Therefore, one's perception of intrinsic motivation is an outcome of the individual's interpretation of the extrinsic consequence, that is, whether the extrinsic consequence is viewed as the reason for learning or as an indication of the sense of competence. Research on this theoretical emphasis has concentrated on the role of extrinsic consequences for learning, including tangible rewards and praise.

Home Environment

Home environment comprises distal and proximal variables. Distal variables, such as socioeconomic status (SES), provide an index of a family's relative demographic position but no direct information concerning the quality or quantity of the home environment experiences to which the child is exposed or the nature of environmental processes as they may influence children's development. These latter process aspects of home environment are referred to as proximal variables and include the cognitive, social-emotional, and physical stimulation available to children in the home as well as family interpersonal relationships.

Relationships Between Home Environment and Academic Intrinsic Motivation

Fullerton Longitudinal Study

The role of both distal and proximal home environment variables has been studied in relation to the development of AIM in the Fullerton Longitudinal Study (FLS), a contemporary, ongoing, prospective investigation, which began in 1979 with 130 1-year-olds and their families and has continuously assessed the participants and their families through childhood, adolescence, and currently into early adulthood. The retention rate of participants was substantial, with no less than 80% of the original sample returning at any assessment. AIM was assessed from ages 9 through 17 using the CAIMI, an instrument developed to measure the cognitive discrepancy, competence/mastery, and attribution aspects of AIM. Both distal and proximal home environment variables have been continuously

assessed from infancy through adolescence. Distal variables include SES, parental educational level, and aspects of parental employment. Proximal home environment was measured with standardized instruments encompassing direct observation in the children's homes, standardized inventories, and parental survey. Analyses indicate both short- and long-term relations between these home environment variables and AIM.

Research Findings

Most of the available research findings concentrate on proximal home environment. This is because analyses of socioeconomic status, parental education, and parental occupation found only inconsistent and low relationships between these variables and children's AIM. In fact, in a longitudinal analysis of the role of proximal environment, that is, cognitively stimulating home environment on children's AIM, it was found that when homes were higher on provision of intellectually stimulating activities, children's AIM was significantly higher from childhood through early adolescence, and this occurred above and beyond their family's SES. Therefore, the provision of proximal cognitive stimulation in the home is important for the development of children's AIM, and not the family SES per se.

One of the areas studied in the FLS concerns the impact of parental motivational practices on AIM and academic achievement. When children were age 9, parents' task-intrinsic and task-extrinsic motivation strategies were assessed using the Parental Motivational Practices Scale. Examples of task-intrinsic strategies are encouraging children's persistence in school work and exposing children to new experiences. Examples of task-extrinsic strategies are rewarding children with money or a toy or removing a privilege. Results showed that these parental motivational practices were significant for children's current and subsequent AIM a year later, as well as for subsequent achievement. Specifically, findings supported the prediction that when parents use task-intrinsic motivational strategies, children will have higher AIM, whereas when parents use task-extrinsic strategies, children's AIM will be lower. Furthermore, parents' motivational strategies had an impact on children's AIM and academic achievement a year later. Hence, findings showed that there was a long-term influence of parental motivational practices on children's motivation and achievement. Children's AIM and achievement are facilitated when

parents encourage children using intrinsic motivational strategies, whereas parents' use of extrinsic strategies has adverse effects on AIM and achievement.

Another study concerned the role of cognitively stimulating home environment on children's AIM. This investigation focused on the short- and long-term relationships between proximal home environment and AIM from childhood through early adolescence. Proximal home environment consisted of the amount of cognitive stimulation available in the home (active stimulation, learning opportunities, and intellectual-cultural activities) as measured through direct observation of the home (including parental interview and observation of the social and physical home environment) as well as surveys completed by the parents. Examples of items from the home observation included (a) child has ready access to a library card, and family arranges for the child to go to the library once a month; and (b) family provides lessons or organizational membership to support child's talent. Examples of items from parent survey include availability of musical instruments and magazines, extracurricular lessons, parental expectation of child's achievement, discussions of political and social issues, and interest in cultural activities such as art, music, literature, and museums. In this study, home environment and SES were measured at age 8. AIM was assessed from ages 9 through 13. Results showed that a cognitively stimulating home environment was a significant, positive predictor of AIM through age 13. Furthermore, a cognitively stimulating home environment had continuous effects on later motivation, as it predicted subsequent motivation through its impact on earlier motivation. When SES was added to the analyses, environment continued to independently and significantly affect later motivation. Therefore, when home environment is more cognitively stimulating, children have greater intrinsic motivation for learning, regardless of their SES. These findings clearly indicate that home environment is a potent factor for children's development of AIM and for sustaining AIM through early adolescence. Specific proximal environmental processes within the family setting contribute to the development of AIM beyond the distal variable of SES. Because AIM is positively related to children's academic effectiveness, it is essential that the home be conducive to stimulating AIM.

As part of the FLS research program, two groups of adolescents have been distinguished on the basis of the consistency of their AIM over time: the motivationally gifted and the motivationally at-risk. These groups were formed on the basis of consistent evidence of either extremely high or low levels of AIM throughout adolescence, respectively. These two groups were found to have distinctly different levels of academic competence, with the motivationally gifted showing pervasively high academic effectiveness and the motivationally at-risk evidencing low academic competency. The role of parental motivational practices, as related to these two motivational groups, also was examined. Because prior research indicated differential relationships between task-intrinsic and task-extrinsic parental motivational strategies and children's AIM, the goal of this research was to determine if parents of motivationally gifted and motivationally at-risk children differentially use intrinsic and extrinsic motivational strategies. Results revealed significantly different strategies used by parents of the motivationally gifted versus at-risk groups. When the children were age 9, parents of the motivationally gifted used significantly fewer extrinsic strategies, whereas parents of the motivationally at-risk group used significantly more extrinsic strategies. At age 17, motivationally gifted adolescents perceived their parents as using significantly more intrinsic strategies, with parents' reports being consistent. On the contrary, at age 17, motivationally at-risk adolescents perceived parents as using significantly fewer intrinsic strategies. Fathers reported using significantly fewer intrinsic strategies, and mothers reported significantly more extrinsic strategies. Motivationally at-risk children are likely to receive significantly more extrinsic and fewer intrinsic parent motivational strategies, a pattern expected to result in reduced AIM. On the other hand, the reverse was true for motivationally gifted children whose parents used significantly higher intrinsic and lower extrinsic motivational practices, a pattern expected to facilitate their children's AIM. Hence, motivationally gifted and at-risk students were exposed to differential parenting messages, which had important consequences for their AIM development.

Specific Subject Areas

Using the CAIMI, the role of proximal environment in the development of AIM for specific subject areas has been investigated. This is particularly important in view of a significant developmental trend obtained

across subject areas in this research program. AIM significantly declines across the school years, from the elementary through the high school years, as do virtually all forms of academic motivation. This is a pervasive finding in the literature, as well as in the FLS. Reading and math are fundamental subject areas, and these show significant declines. The developmental decline in intrinsic math motivation is significantly predicted by the decline in math achievement from childhood through adolescence. This alarming finding creates even greater urgency to determine significant pathways between home environment, AIM, and achievement, as home environment may play a role in ameliorating these declining trends.

The research findings described earlier (i.e., that parents' use of motivational strategies significantly influences children's AIM and achievement) were obtained across all subject areas and for school in general as assessed by the CAIMI. Similarly, the finding that a cognitively stimulating home environment significantly influences AIM above and beyond SES was also obtained for all subject areas and school in general on the CAIMI. Because AIM relates to achievement, whatever can be done to increase parental facilitation of motivation is likely to facilitate children's achievement.

Because of the need to enhance children's literacy and reading achievement, the role of home environment in children's reading intrinsic motivation has been investigated within the FLS. Access to literacy environments that encourage reading engagement is essential, and the home environment can be expected to be one such context. Given this framework, research has been conducted to determine the impact of the home environment on children's reading AIM in the FLS. With regard to the availability of the home reading environment, the following proximal environmental variables have been found to be positively related to reading AIM: variety of experience and maternal involvement (15 months); stimulation of academic behavior and variety of stimulation (39 months); educational stimulation (5 years); level of parental aspirations for the child's educational attainment (5 and 7 years); emotional and verbal responsivity of the environment, encouragement of maturity, active stimulation, educational aspirations, developmentally supportive physical environment, encouragement of achievement, encouragement of persistence and curiosity and involvement in reading activities (8 years); and mothers' expectations of,

and importance placed on, the child's achievement (9 years). Furthermore, examination of causal relationships between home environment and reading motivation and achievement revealed the following: (a) The amount of time parents spent reading to children during infancy and the preschool years had a direct, positive impact on reading AIM as well as reading performance; and (b) availability of reading materials (books, newspapers, magazines) directly influenced reading achievement, and this, in turn, influenced reading motivation. Overall, across the research, a stimulating home environment has been found to be positively related to AIM and achievement across subject areas and to have causal relationships with AIM and achievement, indicating its importance as a possible point of intervention.

Others have corroborated the importance of home stimulation for a child's reading AIM. In a study of fourth- and fifth-grade schoolchildren, two groups of readers were identified: avid and non-avid. Avid readers engaged in more leisure reading outside of school than did non-avid readers. Compared to non-avid readers, avid readers were found to have significantly higher pleasure inherent in reading as measured by the reading scale of the CAIMI, and their parents were significantly more engaged in sharing reading activities with them. These activities included going to the library and reading with the child, giving books as gifts, and encouraging the child to read. These findings are important, as they generalize findings across study populations and indicate that parental practices play an important part in the children's practices and their reading AIM.

Autonomy Provision and Parental Beliefs

A major parenting dimension shown to facilitate the development of AIM is the degree of control afforded children in the home. When children have input into decision making, they are being granted personal autonomy. This aspect of home environment is an important factor in academic motivation. The greater the autonomy parents allow their children, the greater is the children's intrinsic motivation and self-determination. For example, in the study described in the previous section regarding parental motivational practices, task-intrinsic strategies included encouraging independence and mastery of schoolwork and having children answer questions on their own. These

strategies allow children to be more autonomously engaged in their schooling, hence supporting the development of AIM. This parental style is to be distinguished from parental provision of consequences that are extrinsic to the learning per se, as these serve to control behavior rather than allow for children's autonomy.

There is agreement among researchers that parents' provision of autonomy-oriented environments is positively related to children's academically related intrinsic and self-determination motivation. Praising children for learning may have positive or adverse effects on intrinsic motivation depending on whether the praise is perceived as promoting a child's sense of autonomy or, conversely, as controlling a child's behavior thereby resulting in the discounting of intrinsic motivation. The role of praise has been given a great deal of study, as it is pervasively used to socialize children. Whereas parental praise may be well intended, its effect on motivation depends on the degree of control versus autonomy that both the message and delivery convey to the child. For example, when children are praised for succeeding at easy rather than difficult and challenging tasks, or when they are compared to others rather than being recognized for their own competencies, their intrinsic motivation is likely to suffer as a result.

Additionally, parental beliefs and expectations regarding children's achievement strivings, goals, and motivation have been shown to be important for the development of AIM, achievement motivation, and academic achievement. Such beliefs may include parents' perceptions of their children's abilities and interests within subject area domains, concepts of gender-role appropriateness of particular subject areas (such as reading and math), the values parents hold regarding the desirability of achievement in academic domains, or parents' own interests in particular areas that they may then communicate to their children. The more positive the parental values, expectations, and views of the child's likelihood of success in academic activities, the greater should be the child's academic motivation and achievement. The role of parental expectations of success and perceptions of their children's interests, abilities, and effort has been extensively studied, including in the FLS, and results converge on finding positive relationships between parents' beliefs and expectations regarding their child's success and the child's academic motivation and achievement.

Intervention Strategies

Based on the data that proximal home environment has positive and significant relationships to AIM as well as academic achievement, and that many of these relationships support causal interpretations (i.e., that parents influence children's motivation and achievement), recommendations for parental interventions can be advanced. Indeed, many such intervention recommendations have been made, and these are summarized herein. It is important to look again at the three theoretical foundations for AIM because according to their principles, they would implicate different intervention strategies. These foundations are cognitive discrepancy, competence/mastery, and attribution.

Interventions that would follow from the cognitive discrepancy perspective focus upon provision of materials and activities that provide adequate and moderate challenge to children's already existing knowledge and motivation in order to stimulate them to grow beyond these. Specific recommendations include being a careful observer of children's interests and level of engagement in tasks so as to determine the appropriate activity or material that would further stimulate the child by providing optimal cognitive discrepancy. Variety of experience is likely to be successful inasmuch as it provides opportunities for children to be exposed to many areas, to pique new curiosities, and to challenge the child appropriately. Parents should gradually increase the challenge level of materials and follow the child's lead.

Interventions that would facilitate the competence/mastery approach emphasize developing a positive sense of competency that emerges from children being successful in interacting with their environment and in achievement situations. Children's experiences as causing positive outcomes in the environment play an important part in the development of competency perceptions. Play is an important activity that can be an important source of competency development. This context allows children to explore their world without imposed limits. Play provides an opportunity for children to experience feedback from the environment in a setting free from evaluation. Parents should also interact with children to provide them with positive responsiveness.

The attribution approach is most likely to be aligned with the promotion of autonomy perceptions. To the extent that external contingencies, including praise, serve to promote a perception that the child is

working for the reward, or is being controlled by the reward, AIM is likely to be reduced. Conversely, to the extent that parents facilitate autonomy strivings of their children, AIM is likely to be enhanced. Whereas parents may often mean well when they provide children with rewards for achievement, and certainly schools emphasize this type of incentive, this strategy may backfire because perception of the controlling aspects of rewards reduce AIM. Hence, when used, external contingencies and praise should emphasize children's competencies rather than engagement in the activity for the purpose of receiving the reward. Allowing children choice in the selection of learning activities would enhance attributions of autonomy rather than control, and should be encouraged by parents.

Finally, interventions should be tailored to specific subject areas of motivation that are particularly strong or weak in a student. Using an instrument such as the CAIMI can provide a basis for this determination by observing patterns of motivation across subject areas and school in general. Particular strategies may help to support already strong areas and develop relatively weak areas.

Adele Eskeles Gottfried

See also Family Influences; Intelligence and Intellectual Development; Learning; Motivation; Parental Expectations

Further Readings

Gottfried, A. E. (1986). *Children's Academic Intrinsic Motivation Inventory*. Lutz, FL: Psychological Assessment Resources. (Available from Psychological Assessment Resources: http://www.parinc.com)

Gottfried, A. E., Fleming, J. S., & Gottfried, A. W. (1994). Role of parental motivational practices in children's academic intrinsic motivation and achievement. *Journal of Educational Psychology*, 86, 104–113.

Gottfried, A. E., Fleming, J. S., & Gottfried, A. W. (1998). Role of cognitively stimulating home environment in children's academic intrinsic motivation: A longitudinal study. *Child Development*, 69, 1448–1460.

Gottfried, A. E., & Gottfried, A. W. (2004). Toward the development of a conceptualization of gifted motivation. *Gifted Child Quarterly*, 48, 121–132.

Gottfried, A. E., Marcoulides, G. A., Gottfried, A. W., Oliver, P., Guerin, D. (2007). Multivariate latent change modeling of developmental decline in academic intrinsic math motivation: Childhood through adolescence. *International Journal of Behavioral Development*, 31, 317–327.

Gottfried, A. W., & Gottfried, A. E. (1984). Home environment and cognitive development in young children of middle-socioeconomic-status families. In A. W. Gottfried (Ed.), *Home environment and early cognitive development* (pp. 57–116). Orlando, FL: Academic Press.

Gottfried, A. W., Gottfried, A. E., Cook, C., & Morris, P. (2005). Educational characteristics of adolescents with gifted academic intrinsic motivation: A longitudinal study from school entry through early adulthood. *Gifted Child Quarterly*, 49, 172–186.

Gottfried, A. W., Gottfried, A. E., & Guerin, D. W. (2006). A long-term investigation of intellectual and motivational giftedness. *Journal for the Education of the Gifted*, 29, 430–450.

Pomerantz, E. M., Grolnick, W. S., & Price, C. E. (2005). The role of parents in how children approach achievement: A dynamic process perspective. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 259–278). New York: Guilford Press.

Shapiro, J., & Whitney, P. (1997). Factors involved in the leisure reading of upper elementary school students. *Reading Psychology*, 18, 343–370.

HOMELESS FAMILIES

Families with children are the fastest-growing homeless population, although statistics vary between one-half million and a million homeless families in the United States on any given night. Homeless families currently account for about 40% of the homeless population. Homelessness is on the rise in industrialized nations around the world. In this entry, discussion of family homelessness is centered on the experiences of homeless families in the United States, which are somewhat generalizable to other industrialized nations but are inapplicable to nonindustrialized nations torn by wars, famine, disease, and other catastrophic conditions precedent to homelessness.

Definitions of Homelessness

Definitions of homelessness have changed over time, reflecting public opinion, social values, and economics of time and place. In the early 1900s, skid-row residents, vagrants, and people without social networks and resources were considered homeless—heart and hearth were considered together; after World War II, "acceptable" quality of housing became the standard, moving toward definitions of homelessness that equate

it with houselessness. By the 1960s definitions had changed to attribute homelessness to people with no fixed address. Homelessness emerged as a major social problem during the 1980s, under the Reagan administration. In the 1980s, the definition shifted again so that people were considered homeless if they had no private sleeping quarters, regardless of whether that sleeping place changed from day to day.

The Stewart B. McKinney Homeless Assistance Act of 1987, codified in 42 United States Code Annotated (USCA) Section 11302(a), defined a homeless person as an individual who lacks a fixed, regular, and adequate residence or a person who resides in a supervised residence, shelter, welfare hotel, transitional program or place not ordinarily used as regular sleeping accommodations, such as streets, cars, movie theaters, or abandoned buildings. In addition, individuals who are staying in their own or someone else's home but will be asked to leave within the next month were considered homeless. People in jail were not homeless, regardless of whether they would become homeless upon release.

In 2000, the McKinney Homeless Assistance Act was renamed the McKinney-Vento Homeless Assistance Act. Title VII of the McKinney-Vento Act authorized four programs: the Adult Education for the Homeless Program and the Education of Homeless Children and Youth Program, both administered by the Department of Education; the Job Training for the Homeless Demonstration Program, administered by the Department of Labor; and the Emergency Community Services Homeless Grant Program, administered by the Department of Health and Human Services. Since 1995, overall funding for McKinney-Vento programs has declined considerably and the Job Training for the Homeless program has been terminated.

The McKinney-Vento Act does not specifically define homeless families, but families living in shelters and welfare hotels are generally considered homeless while those living in federally or locally subsidized housing are not. Under McKinney-Vento, homeless children and youths are individuals who lack a fixed, regular nighttime residence or who share the housing of other persons because of loss of housing, economic hardship, or a similar reason; are living in motels, hotels, trailer parks, or camping grounds because of the lack of alternative adequate accommodations; are living in emergency or transitional shelters; are abandoned in hospitals; or are awaiting foster care placement. In 2001, Congress again reauthorized the McKinney Education of Homeless Children and Youth Program as the McKinney-Vento Homeless Education Assistance Improvements Act in the No Child Left Behind Act of 2002. McKinney-Vento is equal access legislation; its purpose is to assure every homeless child or youth receives the same educational opportunities, including public preschool education, as other, nonhomeless children.

In practice, homeless families are generally defined as one or more adults accompanied by one or more children under the age of 18. Unaccompanied youths, such as runaways and others separated from an immediate family unit (e.g., children living with family or friends while other family members seek shelter) are not generally included in definitions or counts of homeless families. The 2002 U.S. Conference of Mayors on homelessness in 25 major cities concluded that 41% of individuals that were homeless on any given night were members of homeless families.

Definitions of homelessness that take place within a value system driven by economics tend to include families with children among the "worthy homeless" when determining who deserves aid and assistance. Migrant children who are living in one of the described situations are also considered homeless under the McKinney-Vento Act if they also meet the conditions outlined in section 1309 of the Elementary and Secondary Education Act of 1965. Yet, inconsistent definitions of homelessness and diverse methods for counting homeless people explain discrepancies in reports of the number of homeless people and families nationally and locally.

Causes of Homelessness

The complexity of the problem of homelessness makes it difficult to identify its causes. It is also often difficult to separate causes from effects. In the United States, major natural disasters such as hurricanes in the southern coastal states have suddenly increased the numbers of homeless families. Governmental assistance included efforts to relocate displaced families out of shelters, in part so that shelters would still be available to families. It is too early to know yet whether individuals and families displaced by these natural disasters will be significantly represented in populations of chronically homeless people. Foster care history, drug and alcohol abuse, family violence, runaway, incarceration, and mental and

physical illness can precede or follow a family's experiences of homelessness. Each individual situation needs to be assessed to identify causes or offer remedies. Even so, there are some obvious, systemic causes of homelessness: lack of affordable housing, lagging incomes, domestic violence, and reduced governmental services and assistance.

Lack of Affordable Housing

Researchers agree the primary cause of homelessness is a lack of housing that very low-income people can afford. The Department of Housing and Urban Development sets the standard of "affordable" at no more than 30% of family income. Millions of low-income families pay over half of their incomes for housing or live in severely substandard structures. The shortage of affordable housing began in the 1970s and continues to increase. Related is the problem of delayed returns of children out of foster care because of poor quality housing. The provision of stable housing has been the only measure ever proved to be a solution to family homelessness.

Lagging Incomes

According to the National Alliance to End Homelessness, in no jurisdiction in the United States does a minimum wage job provide enough income for a household to afford the rent for a modest apartment. Federal minimum wage was established as law in 1938. The period from 1997 to 2006 is the longest phase in its history during which minimum wage had not been adjusted. During this time, states set their own minimum wages, higher than those established by the federal government. Two states that report high numbers of hungry people have responded by leading the nation in state minimum wage increases: Washington and Oregon.

Domestic Violence

Violence against women is a widely recognized cause for family homelessness in all regions of the United States. Various studies indicate that about 50% of homeless women with children report domestic violence as the immediate cause of their homelessness. Evictions of women who experience and report abuse are a cause of homelessness of single-mother families. Close to 100% of homeless mothers report severe physical and/or sexual abuse during childhood or adulthood.

Reduced Governmental Services and Assistance

More and varied types of housing subsidies are widely recognized as the single most cost-effective measure government can take to solve the problem of family homelessness. The Personal Responsibility and Work Opportunity Reconciliation Act of 1996 (PRWORA) terminated a 61-year history of entitlement of cash assistance to families. The 1996 revisions to welfare converted Aid for Dependent Children Emergency Assistance and work programs to a block grant, the Temporary Assistance for Needy Families Program, with essentially fixed funding, and it initiated a 5-year-maximum lifetime limit on assistance. PRWORA also reduced food stamp funding by $27.7 billion, and it denied food stamp assistance to legal immigrants. Adjustments were made in 2002, removing some of the barriers for homeless people and legal immigrants and giving states new flexibility in their plans for distributions of food stamps.

Consequences and Conditions of Homelessness for Families With Children

Females in their late 20s with two children under the age of 6 head the majority of homeless families. There are many variations on this pattern of homeless families, however, and the population of homeless school-age children continues to grow. There are single-father heads of households, grandparent heads of households, and two-parent homeless families; all of the myriad variations of family makeup would benefit from affordable housing, minimum wage adjustments, and improvements in the systems of government services to poor and homeless individuals and families. Areas in which the experiences of homelessness have distinct, detrimental impact to people and families, are discussed in this section.

Health

In comparison with the general population, homeless adult family members, primarily females, experience a higher incidence of asthma, anemia, ulcers, upper

respiratory infections, skin diseases, and trauma-related problems and injuries. Homeless mothers are at significantly higher risk for HIV than are low-income housed mothers. Alcohol and drug abuse and dependency, tobacco use, and obesity are also significant health risk factors for homeless mothers.

Asthma, ear and respiratory infections, diarrhea, infectious and communicable diseases frequent to shared living conditions, skin ailments, lice and scabies, lead exposure, delayed immunizations, and poor nutrition and hunger are significant health risks for homeless children, sheltered and unsheltered. Chronic and frequent illnesses of homeless children may prohibit access to services and school attendance. There are also reports that homeless children are at high risk for iron deficiency, delayed growth, and obesity. High levels of stress are thought to cause psychological trauma, depression, anxiety, developmental delays, and behavior problems in homeless children.

Food Security and Hunger

Traditionally, the U.S. Department of Agriculture (USDA) used two labels for food insecurity: "with hunger" and "without hunger." In 2006, the agency renamed food insecurity without hunger as "low food security" and food insecurity with hunger as "very low food security"—leaving hunger out of the definition. Low food security is defined as reports of reduced quality, variety, or desirability of diet, with little or no indication of reduced food intake; very low food security is defined as reports of multiple indications of disrupted eating patterns and reduced food intake. USDA rational for the change is based on the claim that *hunger* is not a scientifically accurate term for understanding food security. The change was further prompted by the USDA's assertion that hunger is an individual, not a household, experience. Critics believe the term *hunger* has political implications, whereas the term *food insecurity* does not. The consequence of renaming food security and of not assessing hunger is still to be determined. Predictions are that more children will be hungry.

Disruption and Dissolution of Families

Homeless youths sometimes become homeless because shelters discourage or do not serve families with older children, especially male children. Families

who are homeless are sometimes able to find housing for their children with relatives and friends, while the parent(s) live in shelters or on the streets. Children living with friends and relatives often have the most profound experiences of instability, caused by separation from parents and other family members; frequent moves between the residencies of caretakers; and occasional stays in shelters when the informal system of housing fails. Neglect, abuse, and family conflict are said to account for the more than 1 million 12- to 17-year-old minors who experience homelessness on their own. One in five homeless children is placed in foster care, and foster care placements in families that have experienced or are experiencing homelessness are typically longer than placements of nonhomeless children. Childhood homelessness is also a precursor of adult homelessness, perpetuating the separations of families intergenerationally. Because family homelessness has only recently been recognized as a major social problem, there is not a substantial research base with which to understand the generational impact of homelessness. Domestic violence is often the most immediate cause for family separations. Reunification of families disrupted by homelessness is rare.

Education and Homeless Families With Children

The relationship between education and family homelessness is particularly complex: Education of parents does not add to the predictive power of who will become homeless. Almost half of homeless adults had not obtained a high school diploma by age 18, and well over half of homeless adults left school altogether or had left school for a period of time during elementary, middle, or high school. There is also evidence that homeless and other poor students are more frequently expelled from schools. On the other hand, well over half of homeless adults do have a high school diploma.

Barriers to school success include lack of access to educational services. Head Start and other publicly funded preschool programs, gifted and talented programs, advance placement programs, special education, family literacy programs, and programs for English language learners have been identified as presenting especially difficult barriers to access for homeless families with children.

Title IX of the No Child Left Behind (NCLB) Act requires, among other things, that the district notify

parents in homeless situations of their child's educational rights. General guidelines are provided to district personnel responsible for coordinating how homeless students are registered, enrolled, transported, linked to community resources, and integrated rather than segregated in classroom environments. Homeless students must be allowed to enroll in school and attend classes even if they do not have all of the required medical records and proof of residency. To stabilize students disrupted from their home settings, the NCLB Act requires schools to provide transportation from shelter to school in some situations. The importance of school as a place of shelter, stability, structure, and accomplishment is sometimes undermined by systems and individual actions that inspire children to report feeling like outcasts, invisible and unwanted. Segregation of homeless children in special schools or classrooms is not allowed under the NCLB Act, except for short periods of time. Exceptions were also made for states already segregating homeless students prior to 2001 (e.g., Arizona). Private foundations have financed some segregated schools for homeless children. Because homelessness disproportionately affects children of minority races and children who speak languages other than English as their primary language, segregation of children living in severe poverty is a controversial affront to efforts to ensure educational equity to all children.

Justifications offered for segregation is that children who are homeless have lower academic achievement and irregular attendance, and experience more emotional, behavior, and mental health issues than housed children. Researchers have demonstrated that homelessness is unlikely to result in long-term depression of children's cognitive or motor development; however, lack of access to education will have such deleterious impact.

Susan Finley

See also Domestic Violence; Family Influences; Poverty; Risk Factors and Development

Further Readings

Finley, S. (Author/Editor). (2006). *At home at school (AHAS) toolkit* [Digital video/DVD-ROM]. Gaston, OR: Kwamba Productions.

Levinson, D. (Ed.). (2004). *Encyclopedia of homelessness* (2 vols.). Thousand Oaks, CA: Sage.

National Alliance to End Homelessness. *Policy focus area: Families.* Retrieved December 2, 2006, from http://www.endhomelessness.org/section/policy/focusareas/families

National Law Center on Homelessness and Poverty. (2007). *Educating homeless children and youth: The guide to their rights.* Retrieved August 25, 2007, from http://www.nlchp.org/view_report.cfm?id=167

Web Sites

National Resource Center on Homelessness and Mental Illness: http://www.nrchmi.samhsa.gov

HOMEWORK

At its basic level, *homework* consists of tasks assigned to students by school teachers that are meant to be carried out during nonschool hours. In most schools across the nation, the daily assignment of homework is as predictable an event as the afternoon dismissal bell. Yet this deeply rooted and quite ordinary aspect of children's schooling is fraught with controversy over its purpose and benefits. Whereas teachers, parents, and even students agree that homework should play some role in schooling, both the amount assigned and the time students should be expected to devote to schoolwork after school remain vexing issues. Research on the extent to which homework boosts academic achievement has yielded contradictory findings, and this has served to fuel the recent anti-homework movement of the 1990s. Fortunately, methodological advances in data analysis, as well as theoretical advances in social cognition, have allowed for a more nuanced and sophisticated understanding of the effects homework on students' learning and motivation.

Studying the impact of homework on academic achievement may seem relatively straightforward, but it is, in fact, multifaceted. According to the model proposed by Harris Cooper, a prominent researcher of homework, the effectiveness of homework varies as a function of (a) exogenous variables (student grade, subject matter, motivation), (b) assignment characteristics (e.g., amount and purpose of homework), (c) classroom factors (e.g., available resources), (d) home-community factors (e.g., whether students have space and materials), and (e) classroom follow-up factors (e.g., teacher feedback). The effects of homework on achievement can be measured

through grades received or on the basis of whether an assignment is completed at all. Such outcomes can be positive or negative, assessed in the short or long term, and can include nonacademic factors, such as whether students are developing more efficient study skills or are experiencing less leisure time. Viewed in this light, the study of homework emerges as rather complex, and the research literature reflects this complexity.

This entry provides an overview of the current state of researchers' knowledge on the benefits of homework, both academically and motivationally. In addition, this entry describes the historical and contemporary context around the practice of homework, considers the relationship between homework and school achievement, addresses motivation in learning—and more specifically, the literature on the relationship between children's beliefs about learning and their school achievement—and examines research on the relationship between homework and motivation.

Perceptions of Homework, Then and Now

The practice of homework went relatively unchallenged for most of the 19th century. With the professionalization of the child study and child health movements in the late 19th century, experts argued that homework, with its focus on drill, memorization, and recitation, was compromising both mental and physical health. In particular, parents and progressive educators were united in the view that homework deprived children of valuable time for play and other worthy extracurricular activities. Indeed, relatively few high school students (8%) reported having 2 or more hours of nightly homework.

With the launch of the Russian spaceship *Sputnik* in the 1950s, serious concerns about American underachievement heralded the advent of the academic excellence movement. Throughout the 1960s and 1970s, homework was seen as a central means to overcome American students' perceived academic deficiencies. By the 1980s, alarming levels of underachievement in American schoolchildren became apparent in both national and international studies of academic performance, and educators continued to emphasize homework as central to fostering academic achievement.

As the 20th century came to a close, widespread media coverage of parental complaints about homework's intrusion on family time prompted renewed debate on homework's effectiveness. Some educators and parent groups seized on contradictory findings of homework's benefits at the elementary school level and argued that the practice should be abandoned altogether. Anecdotal stories of exhausted children, frustrated parents, and beleaguered teachers appeared with some frequency in the print and television media. Interestingly, however, these reports belied national survey data and empirical findings on perceptions of homework. For example, a national survey of parents whose children attended public schools revealed that only 10% believed their children were assigned too much homework. Even if they have difficulty helping their children, most parents agree that this is a natural and expected aspect of parenting and believe that homework helps to advance their children's learning. Teachers welcome parental involvement because it fosters both home–school connections and school success. At all ages, students acknowledge the role that homework plays in helping them learn, a belief that is particularly strong at the high school level. By this time, parents, students, and teachers agree that there is a relationship between completing homework assignments and school achievement. Indeed, more than 75% of teenagers believe that they would learn more if their teachers enforced homework policies.

Data collected as part of the ongoing studies of the National Assessment of Educational Progress, conducted at regular intervals by the U.S. Department of Education, reveals information regarding how much time students spend on homework and how this varies as a function of age. While the percentage of 9-, 13-, and 17-year-olds who report having no homework has decreased since the 1980s, 20% of 9- and 13-year-olds and more than 25% of 17-year-olds in 2004 affirmed that they are assigned no homework at all. At all three age levels, the proportion reporting having less than 1 hour of homework has increased since the 1980s (from 41% to 59% of 9-year-olds, 32% to 40% of 13-year-olds, and from 24% to 28% of 17-year-olds). At the same time, there has been virtually no increase in the proportion of students at any age reporting either 1 to 2, or more than 2 hours of homework. In fact, reports of more than 2 hours of nightly homework are at their lowest since such data have been available (4% of 9-year-olds, 8% of 13-year-olds, and 11% of 17-year-olds). Although these figures may not suggest that American schoolchildren are overly burdened with homework, it

is important to examine the relationship between homework and school achievement.

Relationship Between Homework and Academic Achievement

Overall, there is a strong and positive relationship between homework completion and academic achievement. In general, students who do more homework have higher grades, even controlling for prior grades, demographic background, and a variety of after-school activities. This relationship is particularly strong at the secondary school level. Among younger children (second through fourth grades), however, there tends to be a negative relationship between the amount of time spent on homework and academic achievement. As mentioned earlier, this finding has led some parents and educators to argue that homework should not be assigned to younger children.

Why should homework benefit older, but not younger, children? It appears that teachers adhere to different beliefs about the value of homework for students at different ages. Elementary and secondary school teachers alike report that they assign homework to help students learn and to foster study and time management skills. Relative to secondary school teachers, however, elementary school teachers place greater emphasis on homework's value in fostering time management. This suggests that, in the younger grades, homework may be viewed as more useful for the development of skills that will pay off in the long run, and its effects on school grades may not necessarily be observed in the short run. Additionally, younger children may be less skilled at ignoring distractions and focusing their attention on their homework when they are at home, and they may also have less effective study habits. Thus, at younger ages, the benefits of homework may be "in the making" and nonacademic in nature.

Responding to concerns over the lack of clarity between homework and academic achievement, researchers have suggested that methodological weaknesses and problematic operational definitions of achievement and homework have contributed to contradictory findings on homework's effectiveness. For example, many experimental and quasi-experimental studies, in which students or classes are assigned to homework or no-homework conditions, have not employed randomization techniques. Furthermore, it is problematic to compare grades across classes when teachers differ in the amount of homework they assign. Finally, time on homework, a major variable in the research, is determined differently by different students, depending, for example, on how motivated or conscientious students are.

Of course, the demonstrated associations between homework and achievement outcomes do not speak to the issue of causality. Does homework foster academic achievement, or do stronger, more motivated students choose to do their homework? In a meta-analysis (an analysis of many studies on this topic), Cooper and his colleagues examined research conducted over a 16-year period in which investigators tested the causal relationship between homework and achievement. Their analysis included studies that used four different research designs: (1) random assignment of classrooms, or students within classrooms, to homework and no-homework conditions; (2) naturalistic, cross-sectional studies of classrooms in which homework was assigned; (3) naturalistic studies of the amount of homework students do and its relationship to achievement outcomes, controlling for other influential variables; and (4) structural equation models that used naturalistic data from the High School and Beyond national database.

Most of the included studies found a significant causal relationship between the amount of homework students report completing and their school achievement. Cooper noted, however, that each of these studies was problematic in some way, which calls into question the claim of causality. At the same time, he noted that this wide range of studies was not flawed in the *same* way, rendering it likely that the studies do support a causal relationship between homework and achievement. In other words, multiple empirical studies, conducted through diverse research designs and across multiple school subjects, have found a strong and positive relationship between the amount of homework students do and their school achievement.

In addition to its apparent academic benefit, homework also serves a motivational benefit, in the sense that it has the potential to foster adaptive skills that become more critical as children advance into their middle and high school years.

Motivational Factors in Learning

Achievement motivation refers to goal-directed beliefs and behaviors that children display in educational settings. These include both cognitive and affective factors, such as self-perceptions of ability, attributions for

success and failure, and academic self-worth. These motivational factors influence both the goals that students choose to pursue and the persistence with which they pursue them. For example, children who believe that intelligence can grow with effort tend to embrace challenge and view mistakes as opportunities to learn. In contrast, children who view intelligence as more of a static entity tend to avoid challenge and perceive mistakes as condemnations of their abilities. Bernard Weiner, whose prominent theory of achievement motivation has shaped much of the contemporary research in achievement motivation, has demonstrated that students' attributions for success and failure predict their affective reactions to their performance in school, which in turn predict their achievement behavior. For example, if students believe that poor performance on a test is due to lack of ability (a quality largely perceived as internal, stable, and uncontrollable), they are less likely to invest future effort in their work. In contrast, if they believe that a failure was due to lack of effort (generally perceived to be internal, unstable, and therefore controllable), they are more likely to renew their future efforts.

How do children develop beliefs about learning? Not surprisingly, parents have a profound influence on their children's growing awareness of their abilities. Researchers agree that parents' beliefs about learning serve to guide their educational socialization practices, and their beliefs and practices have predictive impact on their children's beliefs. For example, a longitudinal study of self-perceptions of competence in more than 300 children, beginning in kindergarten or first grade until the fifth grade, showed that, by the fifth grade, parents' perceptions of their children's competence strongly predicted children's self-perceptions of their math ability.

Teachers, in their roles as facilitators of learning, have enormous influence on children's developing achievement beliefs. The literature on expectancy effects has shown that low teacher expectations serve to depress children's self-perceptions of ability and their achievement outcomes. As institutions, schools also serve to communicate their beliefs about students' ability to learn, through practices such as ability grouping at the elementary level and tracking at the secondary level. Relative to students in higher tracks, those in lower tracks report learning in classrooms where there is much disruption, more punitive teachers, and curricula oriented around rote memorization and basic skills. Not surprisingly, lower-tracked students tend to have lower self-perceptions of ability and diminished aspirations for the future.

As an educational practice, homework provides opportunities for parents and teachers to communicate positive beliefs and self-perceptions about achievement and the nature of ability; in this way, homework serves as a training ground for the development of adaptive beliefs about learning. Resistance to this view exists, despite research evidence to the contrary.

Homework and the Development of Motivational Skills

The public discourse over the motivational advantages of homework has been as contentious as the debate over its academic benefits. Critics of the practice argue that homework robs children of time to play, imposes an unfair burden on parents, and fosters stress in children and their families. Furthermore, critics argue, there is no research evidence that supports the view that homework fosters motivational skills, such as more effective studying, better organization, and self-regulation. Interestingly, though, most parents believe that the academic (homework help) and nonacademic (scaffolding organization and self-regulation) tasks they perform on a daily basis constitute important aspects of their role as parents. Parents' lay theories of school success are confirmed by research evidence that shows that homework can influence the development of motivational skills.

Parental involvement itself has been conceptualized as multidimensional, with parents demonstrating their involvement behaviorally (attending school events), personally (showing enjoyment of children's school), and cognitively and intellectually (providing help). Homework serves as a means for parents and teachers to set high standards and improve performance, and it is an important variable mediating the relationship between high academic motivation (including school interest and aspirations) and school achievement. Involvement in homework, then, gives parents opportunities to foster critical motivational and behavioral qualities in their children, such as self-regulation, perceptions of self-efficacy, school engagement, conscientiousness, and mastery orientation, all of which, in turn, are associated with school success.

Investigators recently demonstrated that parents who adopt a mastery-oriented approach when helping their children with homework can enhance their children's psychosocial well-being, even among those with

negative self-perceptions of ability. The researchers studied parents' reported homework assistance patterns (e.g., focus on understanding material, support for autonomy) and monitored children's (third, fourth, and sixth graders) perceived competence, mastery orientation, and positive and negative emotional functioning. Mothers' mastery-oriented practices predicted enhanced perceptions of academic competence, mastery orientation with schoolwork, and positive emotional functioning among those children with more negative perceptions of competence. In other words, mothers who fostered positive beliefs about ability and emphasized the value of effort and the importance of mistakes in learning promoted positive emotions in their children while they were doing their homework and positive self-perceptions of competence.

Implications

The debate over homework is not new, and it has tended to vary as a function of available theories of child development, as well as perceptions of the national state of student achievement. Academic excellence and high-stakes testing are currently playing central roles in children's education. In this context, homework is perceived as vital to academic development. At the same time, however, the average American student does not appear to be overly burdened by homework assignments.

The research literature has established that there is a strong and positive relationship between homework and achievement. Notably, researchers continue to fine-tune ways in which this relationship can be better understood. At the same time, rigorous evidence from the field of achievement motivation demonstrates that the practice of homework serves to foster the development of adaptive approaches to learning, which are increasingly necessary as students progress through to the postsecondary world of advanced study and work.

Janine Bempechat

See also Home Environment and Academic Intrinsic Motivation; Learned Helplessness; Motivation; Parenting; Self-Efficacy; Teaching Strategies

Further Readings

Bempechat, J. (2004). The motivational benefits of homework: A social-cognitive perspective. *Theory into Practice*, *43*(3), 189–196.

Cooper, H., Robinson, J. C., & Patall, E. A. (2006). Does homework improve academic achievement? A synthesis of research. *Review of Educational Research*, *76*(1), 1–62.

Duckworth, A. L., & Seligman, M. E. P. (2005). Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological Science*, *16*(12), 939–944.

Eccles, J. S., Roeser, R., Vida, M., Fredricks, J., & Wigfield, A. (2006). Motivational and achievement pathways through middle childhood. In L. Balter & C. S. Tamis-LeMonda (Eds.), *Child psychology: A handbook of contemporary issues* (2nd ed., pp. 325–355). New York: Psychology Press.

Gill, B. P., & Schlossman, S. (2003). A nation at rest: The American way of homework. *Educational Evaluation and Policy Analysis*, *25*(3), 319–337.

Hoover-Dempsey, K. V., Walker, J. M., Sandler, H. M., Whetsel, D., Green, C. L., Wilkins, A. S., et al. (2005). Why do parents become involved? Research findings and implications. *The Elementary School Journal*, *106*(2), 105–130.

Loveless, T. (2003). *Brown Center report on American education 2003: How well are American students learning?* Washington, DC: Brookings Institution.

Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom.* New York: Holt, Rinehart & Winston.

Trautwein, U., & Koller, O. (2003). The relationship between homework and achievement: Still much of a mystery. *Educational Psychology Review*, *15*(2), 115–145.

Weiner, B. (2005). Motivation from an attributional perspective and the social psychology of perceived competence. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 73–84). New York: Guilford Press.

I

It's not that I'm so smart, it's just that I stay with problems longer.

—Albert Einstein

IDENTITY DEVELOPMENT

One important task in life is the establishment of identity, a task usually associated with the adolescent. As part of Erik Erikson's theory of psychosocial development, he described the major crisis occurring during the teen years as being identity versus identity confusion. *Identity* is the individuality of a person and comprises relatively stable personality traits that are especially prevalent in certain situations (role behaviors such as being a parent). *Identity development*, then, is the search for this person one will become, which can either be a smooth or quite complex process.

According to Erikson, a person must resolve his or her adolescent crisis before the person can confront the next stage crisis; thus, some sort of solution must be made with regard to identity, making it a rather important topic. Identity development is not a simple task because it is a process of understanding and synthesizing childhood experiences and memories, family history, and present self-awareness and acceptance, as well as future goals or ideals. Beyond this internal process, the person must also combine his or her relationships with persons and things external to him or her.

Although a development of identity occurs throughout the life course, most of the focus is placed on the adolescent years. The reason for this focus is that adolescence marks two important developments regarding identity. The first is that this is the time in which a person has his or her first experience with reorganizing and restructuring his or her sense of self. The second is that adolescence also marks the time in which cognitive functioning and intellectual capability are high enough to understand the importance of these changes.

This entry examines many aspects important to identity development, including identity crisis, identity status, negative identity development, gender differences, changes in adolescence affecting identity, ethnic identity, and gender role development.

Identity Crisis

As with the other eight stages of Erikson's psychosocial development model, identity development centers on the term *crisis*. This term does not necessitate a negative conflict, however; it is also synonymous with the term *exploration*. This more neutral term helps to describe a crisis as (a) a developmental turning point, where the person is more exposed but with high potential; and (b) a point in time where the individual is in search of his or her new identity and is trying out options and alternatives.

Erikson argued that crisis is a universal occurrence, although some persons may complete this process in the unconscious and thus do not realize it is happening. One important realization of crisis that must be understood is the fact that it is not a single occurrence that leads to fast resolution, but that it can be gradual and occurring at multiple times across many years.

When one crisis is settled, another can quickly occur. The typical focus of crises centers on occupational and ideological decisions, as these are fundamental to other decision making throughout the life course.

Two types of crises are defined. The first is associated with the adolescent years, especially in males. The other has its focus on persons who already have established identities and typically targets females.

Identity Deficits

Identity deficits occur when an individual rejects familial morals, values, and plans so as to determine the fate of his or her own identity. This type of crisis leads to a large amount of exploration of ideas, lifestyles, and things that are previously unknown (or at least inexperienced) to the individual. This type of crisis can alternate between feelings of excitement to feelings of depression and anxiety in the person, and it is sometimes described as a turbulent time. This type of crisis is typically resolved through the interactions the individual has with important persons in his or her environment and the selection of attributes, values, and opinions of these persons.

Identity Conflicts

Identity conflicts occur when a person realizes that his or her multiple identities and typical role behaviors lead to inconsistencies or conflicts in determining action. This type of conflict is not associated with the same highs as deficits and are usually associated with feelings of depression and guilt. Resolving conflicts typically necessitates a compromise and loss of parts of a person's identity.

Identity Status

Four statuses of identity have been identified and are determined by the initiation and resolution of crisis. Although these are sometimes referred to as stages, it is important to understand that these stages cannot be viewed linearly and thus do not have a hierarchical structure; however, achieved identity is more positively associated with beneficial outcomes.

Achieved Identity

A person is labeled as having an achieved identity if he or she has gone through a crisis and had a resolution that has created commitments (career, relationship, value system). These persons are seen as being more stable and mature.

Moratorium

Moratorium is a time of continued exploration. A crisis has occurred but no resolution has been established, thus no commitments. Moratorium was a large focus of Erikson's work in identity, and he described youth as being stuck in this stage, because although they may have ideas and goals of what they desire in life, they are not yet able to make the commitments or achieve these goals. These persons are described as being unstable and receptive to many new opinions and lifestyles.

Foreclosed Identity

A person with a foreclosed identity has commitments but has not gone through the process of crisis or exploration. This identity may have been imposed on the person by others, or the person may have prematurely determined his or her life course. This type of person is described as being mature, rigid, somewhat shallow, and many times very close to his or her parents.

Identity Diffusion

This stage of identity is absent of both crisis and commitment. This person is seen as being immature and may be prone to psychopathology.

Negative Identity Development

Beyond the previously mentioned identity statuses, Erikson placed much focus on the idea of developing a negative identity. Negative identity occurs when a person selects an identity that is not accepted by the person's family and/or community. This identity is typically established to spark a uniqueness so as to stand out, whereas an accepted identity would only cause the person to be lost in the crowd. This type of identity creates negative attention, but it is believed that this negative attention is better than no attention.

Gender Differences in Identity Development

As mentioned earlier, there are differences across the genders with regard to identity development. Beyond

the different crises associated with each gender, females are more concerned with establishing identities around relationships with persons in their environment, whereas males are more concerned with occupational and ideological identities. Females are also more associated with achieved identities and moratorium (centering on family/career priorities), and males are associated with foreclosed and diffused identities.

Changes in Adolescence Affecting Identity

Major changes occur across adolescents that have a profound effect on identity development. These changes can be viewed while examining changes in a person's self-conception and self-esteem.

Self-Conceptions

Self-conceptions are how a person regards him- or herself in relation to traits and attributes. As stated earlier, when a person enters adolescence, his or her cognitive capabilities increase, which allows the person to think more abstractly and to process greater amounts of information. This leads to a greater ability to define oneself, especially creating multiple definitions that change depending on the situation present. This idea that personality can be multifaceted is a new concept in adolescence and is important for the development of identity.

Self-Esteem

Self-esteem is how positively or negatively a person feels about him- or herself. Despite popular opinion, there is not a dramatic drop in self-esteem during adolescence, although there are fluctuations through the years. This is, however, the time in which self-esteem becomes more stable and changes less in different situations, which leads to a more stable sense of identity.

Ethnic Identity Development

Ethnic identity is typically not a major role of identity development until late adolescence or early adulthood. Although ethnic identity is typically thought of as a minority issue, White youths also develop a sense of ethnic identity, although this is usually substantially weaker than their minority peers. One way White youths strengthen their sense of ethnic identity

is to identify with ancestral ethnic roots (Jewish, Italian, Irish, etc.).

The research examining the establishment of an ethnic identity has developed numerous scales discussing how a particular ethnic group establishes identity. In this entry, a Racial/Cultural Identity Model (R/CID) created by Derald Wing Sue and David Sue is presented that uses similarities between the identity development processes that all ethnic groups share to create a single model explaining development.

In their model, they present five stages through which a person passes in relation to his or her views of the dominant and minority culture: conformity, dissonance and appreciating, resistance and immersion, introspection, and integrative awareness.

Conformity

In this stage, a person is more in line with the views of the dominant culture than the person's own minority culture. This type of person is typically negative toward him- or herself and other members of his or her culture of origin, as well as other minority cultures, while thinking highly of White culture and emulating it.

Dissonance and Appreciating

In this stage, a person currently in the conformity stage comes in contact with information that contradicts his or her views of his or her minority culture. Although this encounter does not produce a direct change, it begins to modify the person's beliefs, and a conflict begins to erupt as the person questions stereotypes he or she has held of his or her own minority culture and other minority cultures.

Resistance and Immersion

Once the person has been in the dissonance and appreciating phase for a while and has had experiences denying his or her original beliefs, the person begins to completely reverse these beliefs and adopt his or her minority culture almost completely. This also causes a rejection of the dominant culture in such a way that the person also reacts against it. Due in part to the person's original denial of his or her minority culture, the person in this stage typically feels anger, guilt, and/or shame, with the anger directed at oppression and racism.

Introspection

After some time in the previous stage, the person begins to realize the amount of energy he or she loses feeling angry and reacting toward the dominant culture. The person realizes that he or she loses opportunities to truly learn about him- or herself, and the person also begins to experience times in which his or her negative views of the dominant culture are false. This again leads to a conflict between the person's views of the dominant culture while he or she still holds on to his or her minority culture views.

Integrative Awareness

This is the stage in which a person can appreciate views from his or her own minority culture and those of the dominant culture, and the views are not in conflict or opposing. This person develops strong self-esteem and confidence. He or she is able to have a strong pride in his or her minority culture while obtaining autonomy, as the person appreciates all cultures.

Gender Role Identity Development

A last major portion of a person's identity stems from his or her definitions of appropriate gender roles. All persons differ in the amount of masculine and feminine traits that they exhibit. Throughout life, gender roles are imposed on a person, but this pressure increases as a person enters adolescence, because although the person is becoming more open to multiple role behavior, the pressure to act in stereotypical roles is increasing. A further complication arises for males especially because research has shown that masculine traits are more accepted by peers and lead to greater self-esteem across both genders. This means that females can be more androgynous, exhibiting both masculine and feminine traits, but males must exhibit only a single trait.

Matthew J. Davis

See also Erikson's Theory of Psychosocial Development; Psychosocial Development

Further Readings

Erikson, E. H. (1959). *Identity and the life cycle.* Madison, CT: International Universities Press.

Erikson, E. H. (1968). *Identity: Youth in crisis.* New York: Norton.

Santrock, J. W. (2002). *Life-span development.* New York: McGraw-Hill.

Steinberg, L. (2005). *Adolescence* (7th ed.). New York: McGraw-Hill.

Sue, D. W., & Sue, D. (2003). *Counseling the culturally diverse: Theory and practice* (4th ed.). New York: Wiley.

IMMIGRATION

The current global trend is one of transition and adaptation to new contexts of living. *Immigration* refers to the movement of people from one nation-state to another, with the term *migrant* referring to a person who has settled outside of his or her country of origin. It was previously believed that people tend to move from developing nations to developed nations, or from Third World countries to First World countries. However, the Global Commission on International Migration reports that all countries in the world are currently receiving migrants from neighboring countries or from abroad, and are also experiencing outward migration flows. Statistics from the International Organization for Migration suggest that there are currently 191 million migrants worldwide. Whereas previous waves of immigration to North America were disproportionately European, current immigration patterns show a diversification of migrants' source countries to include Asia and the Pacific, Central and South America, Africa, and the Middle East. Globalization and immigration contribute to the creation of pluralistic societies, where education and health care systems are faced with the challenge of understanding and responding to the needs of multilingual and multicultural constituents with unique life histories. This overview of immigration begins with a description of different migrant subgroups and the common acculturation process they experience in the new country of settlement. The implications of immigration for education and life span human development, as well as for mental health, are subsequently discussed.

Subtypes of Migrants

Migrants can be subdivided into three groups based on whether their relocation outside of their countries of origin is voluntary or involuntary and based on the degree of permanence of their settlement in the host

nation. These three groups are immigrants, sojourners, and refugees. Each of these groups is described below, along with the process of acculturation that they commonly experience as part of their adjustment to a new sociocultural milieu.

Immigrants

Immigrants are people who voluntarily leave their countries of origin in search of better social, educational, and vocational opportunities, with the intention of permanent residence in the destination country. Conditions in their countries of origin, such as overpopulation and high rates of unemployment, or natural disasters that have adversely affected their quality of life, may serve as motivators for immigration, in addition to perceived opportunities abroad. Immigrants can gain entry to their preferred destination country under either of two immigration categories: the family class or the economic class. *Family class immigrants* are those who are sponsored by an immediate family member or relative living in the destination country who makes a commitment to take temporary financial responsibility for them and to assist with settlement and integration. *Economic class immigrants* are those who qualify to independently migrate to the destination country based on having education and training that corresponds to critical labor market needs, or based on investment capital that supports their potential for entrepreneurship.

Sojourners

Sojourners are people who voluntarily leave their countries of origin to temporarily resettle in another country for the purpose of obtaining an education or fulfilling employment responsibilities, or for humanitarian reasons. International students, foreign diplomats, migrant laborers, and disaster recovery personnel, such as medical professionals in the United Nations Doctors Without Borders Program, are examples of sojourners. Sojourners typically gain entry into their countries of temporary settlement based on applications for international student visas, work permits, or permits for foreign aid personnel. Sojourners may be living in a country other than their country of origin for many years. Besides adjusting to life in the country of temporary settlement, when sojourners return home, they often have to reacclimate to their previous way of life.

Refugees

In contrast to immigrants and sojourners, refugees are people who are forced to leave their countries of origin due to sociopolitical circumstances that pose risks to their personal safety. Depending on whether conditions in their countries of origin improve over time permitting a possible return, such as a change in political leadership, refugees may perceive their immigration to be permanent or impermanent. The United Nations definition of refugees highlights their potential for persecution based on their race, religion, nationality, social group membership, or political opinion if they remain in their home countries.

The United Nations' criteria for defining a refugee have recently been expanded to include people fleeing their home countries because of war or civil conflict. Current global statistics indicate that political and social upheaval affects the welfare of people in more than 130 nations, prompting unprecedented levels of forced displacement. Large numbers of refugees remain internally displaced near the borders of their home countries before being able to seek asylum abroad. The International Organization for Migration reports that 23.7 million people are currently internally displaced in or near their countries of origin. Asylum seekers claim refugee status upon or after arrival in their destination country, although some people reside in the destination country illegally and are undocumented. Refugee claimants are granted permanent residence in the country of asylum based on an adjudication process guided by humanitarian considerations.

Acculturation

John Berry identified acculturation as the common experience among migrant groups residing in pluralistic societies. He has also presented the most widely used and most comprehensive framework linking acculturation with policies related to multiculturalism in migrants' destination countries. Berry describes acculturation as the process of sociocultural transition whereby migrants make two related decisions: (1) how much of their unique cultural values and behaviors they will retain or relinquish, and (2) how much they will interact with other cultural groups in their new country of residence. The variable outcomes of these two decisions lead to one of four acculturation strategies: separation, assimilation, integration, or

marginalization. Separation involves cultural maintenance and limited intergroup contact. Assimilation involves immersion into the surrounding society and adoption of the values and behaviors of the dominant group in that society, with limited heritage culture retention. Integration involves combining the culture of one's home country and current host country in values, behaviors, and interaction patterns. Marginalization refers to ambivalence toward both the heritage culture and the surrounding culture.

National and regional policies related to multiculturalism in migrants' destination countries inevitably affect the acculturation strategies available or unavailable to them. For example, during apartheid in South Africa, Black migrants from other regions were segregated and did not have opportunities for intercultural contact or exchange. In contrast, in the United States, the "melting pot" policy encourages assimilation. In countries such as Canada and Australia, where national policies encourage intergroup contact while supporting heritage culture maintenance, migrants have maximal opportunities for integration. Besides national policies, migrants' acculturation strategies also vary depending on whether their immigration was voluntary and whether they perceive their move to be permanent or impermanent.

Immigration, Education, and Human Development

Childhood and Adolescence

Immigrant and refugee children entering the school system in the new host society are faced with the task of mastering the primary language of instruction, which most often is not their native language. English is the primary language of instruction in school systems in North America, as well as the most widely recognized international language. Children's entry level of English proficiency and academic histories in their countries of origin affect their school adjustment and academic performance. War and civil unrest in many countries of the world have caused long-term interruption to the schooling of large numbers of refugee children, who have been displaced in refugee camps for several years. Children who have experienced school interruption at early stages of the learning process do not have the opportunity to acquire basic literacy skills and verbal comprehension skills in their native language, impairing English acquisition. A common strategy for second-language learning is making analogies between words, structures, and meanings in the first and second languages, and refugee children may lack these anchors for second-language acquisition. School-based programs in English as a second language (ESL) sometimes require modification for refugee youth, beginning with preliteracy tasks followed by graded increases in literacy and verbal skills.

Because the majority of cognitive, intellectual, and aptitude tests used with schoolchildren in North America are administered in English, immigrant and refugee children's performance may be compromised by limited English proficiency. Children from these migrant groups have been erroneously identified as having limited intellectual capacity or as being learning disabled and are placed in special education programs, leading to the passage of various legislative acts and bills to promote nondiscriminatory and culturally sensitive assessment practices. Nondiscriminatory assessment practices involve the following:

- Considering and ruling out alternative causes of learning failure in the diagnosis of intellectual or learning disabilities,
- Examining standardization samples to ensure that test norms are applicable to the immigrant/refugee children being tested,
- Evaluating critically a test's cross-cultural reliability and validity, and
- Refraining from administering tests that have been found to result in the disproportionate identification of migrant or minority children as cognitively or intellectually challenged, compared to their nonmigrant peers.

When immigrant/refugee children are able to acquire English as a second language during early periods of development, they may experience advances in cognitive development. Childhood bilingualism has been linked to increased cognitive flexibility and sophistication of thought. It has also been linked to higher academic achievement. Children and youth who are bilingual can also effectively alternate between different languages and modes of communication based on the immediate context of interaction, a phenomenon referred to as *code switching*. Code switching offers a clear communication advantage in pluralistic societies.

Although language acquisition is one key element of the culture learning that occurs when all migrant groups adjust to a new society, children and adolescents also

learn more about their racial/ethnic and cultural identities through interactions with peers in host society schools. Ethnic minority and migrant children generally show earlier levels of racial/cultural consciousness than do children from the dominant group in pluralistic societies because of their distinctiveness from the dominant majority. In adolescence, identity development becomes particularly salient. In Erik Erikson's theory of psychosocial development, the key task of the developmental stage of adolescence is the formation of a consolidated personal identity. Achievement of an identity involves exploration of a variety of values and behaviors, with the result of making a conscious commitment to a chosen way of being. For migrant youth, establishing a consolidated identity involves examining and resolving their attitudes toward their own cultural group and the dominant group in the new host nation, in direct parallel with the acculturation process. Interference with identity development can occur when experiences such as racism and discrimination impair youth's attempts to integrate with and explore host group ways of life, or when parents' emphasis on exclusive cultural maintenance precludes such exploration.

When children or adolescents' school experiences and identity exploration process produce changes in behavior that depart from cultural norms, familial conflict or parenting stress may result. In Erikson's theory of psychosocial development, the key developmental task of midlife is to make a valuable contribution to the lives of future generations. The midlife stage in Erikson's theory represents the age bracket of most parents of adolescents, for whom this task translates into an attempt to provide their children with the skills needed to become successful members of society. For immigrant and refugee parents, a central part of the parenting process is the transmission of their cultural heritage. Youth's behavioral shifts away from cultural norms may therefore prompt parental efforts to redirect children to traditional behavior patterns. School counselors are often placed in the position of "cultural brokers" who assist parents and youth to manage the family's cultural transition process.

The cultural transition process and identity development process pose unique challenges for international students in both secondary schools and college/university settings. Second-language acquisition and some degree of cultural change facilitate their adaptation in the host country. However, many international students experience "reverse culture shock" upon returning to their country of origin, finding that they can no longer identify with their previous social network because of the self-changes that occurred abroad. Furthermore, changes in population density, employment opportunities, and social or political conditions in their countries of origin may result in a stark difference between their life in the predeparture phase and their reentry experience. Although international student programs and centers in academic settings prepare them for their tenure in the host country, counselors in these settings attempt to also prepare them for a successful transition home. International students reflect globalization in its truest sense, often maintaining transnational ties and becoming involved in multinational corporations after completion of their education.

Adulthood

Migrant adults often experience challenges integrating into the North American labor market. Some attribute challenges to factors such as language barriers, racism, and discrimination. However, the most pervasive barrier to participation in the North American labor market relates to their foreign educational credentials. Although there is significant variability in levels of education among migrants entering a given country, immigrants entering the destination country through the economic class as well as some groups of refugees are highly educated. Because economic class immigrants are selected for immigration based on their educational and professional credentials, they expect to obtain employment in their preimmigration occupations after relocation. Many are dismayed to find that their foreign educational qualifications are not recognized by professional associations governing their occupations in their new host country, or by employers. The common and widespread experiences of unemployment and underemployment among highly educated migrants contribute to the downward socioeconomic mobility of families after immigration.

Difficulties with obtaining employment after immigration have resulted in a trend toward adult education among immigrants and refugees. Many retrain to enter their professions in the new country or seek education or apprenticeship opportunities to consider alternative career paths. Continuing education and career development in the new country present opportunities for upward social mobility and eventual recovery of the family's preimmigration socioeconomic status. Migrant

parents may also place a great deal of emphasis on the education of their children, as children may be viewed as key agents in the continued upward mobility of the family.

Immigration and Mental Health

Early studies comparing the mental health status of migrant and nonmigrant residents in various North American societies found that migrants have significantly poorer mental health and higher levels of psychiatric symptoms than nonmigrants. These findings were used to support the social drift theory of mental illness. The theory suggests that when migrants cross international borders, they bring with them multiple pathologies, thereby lowering overall indices of population health status in a given geographic region. In the past few decades, new research studies have revealed that differences in the mental health status of migrant and nonmigrant populations only appear among psychiatric in-patient samples, with a reverse pattern of findings among community samples. The consensus in the present research is that there is no inherent difference in the mental health status of migrant and nonmigrant groups. However, it is generally acknowledged that the multiple stressors that migrants may encounter at various stages of the immigration process may make them vulnerable to mental health problems if the stressors exceed their coping resources (material assets, level of social and instrumental support, and personal coping skills). The preimmigration and postimmigration circumstances of each migrant group may pose unique challenges to members' mental health and adjustment. The following sections identify factors related to the mental health of each group and consider protective factors that may promote resilience in the face of adversity.

Immigrants

As described in the previous sections, immigrants face a number of adjustment challenges in the new country, including learning a new language, seeking education and employment, establishing a bicultural identity, dealing with experiences of racism and discrimination, and dealing with family conflicts related to the cultural transition process. All of these factors interact in influencing their mental health status. For example, besides possible conflicts in the parent-child subsystem after immigration, challenges in obtaining

employment among adult family members often produce changes in the husband-wife subsystem. Immigrant women tend to obtain employment more rapidly than immigrant men. After immigration, women who remained at home with their children in their countries of origin often need to enter the workforce to ensure the family's financial survival. Women's entry into the labor market and the transformation of single-income families to dual-income families often results in changes in marital roles and relationships. Role changes may lead to family conflict, reducing family cohesion and increasing stress levels, thereby contributing to mental health problems.

Refugees

In addition to dealing with the potential stressors that immigrants face upon their arrival in the host society, millions of refugees have experienced various forms of torture or other traumatic events prior to migration or during exodus from their homelands. Richard Mollica is credited with identifying subtypes of adverse events that many refugees have experienced before their arrival in the country of asylum through his development of the Harvard Trauma Questionnaire. These include the following:

- deprivation of food, shelter, or interpersonal contact;
- physical injury, including sexual assault and torture;
- incarceration and forced participation in re-education camps; and
- witnessing the torture or killing of other people.

The World Medical Association describes torture as the act of inflicting physical or mental suffering on another human being for reasons such as obtaining information or eliciting confessions. The Association suggests that torture is both deliberate and systematic, and it can be inflicted by an individual or group acting on the orders of an authority figure.

Refugees have to concurrently deal with any human rights violations that occurred in their countries of origin; the possible breakdown of their family, social, and cultural support systems; and the burden of resettlement in a new society. Common health problems stemming from these life experiences include posttraumatic stress disorder (PTSD) and other anxiety disorders, depression, and multiple physical injuries and ailments. Symptoms of PTSD among refugee groups have been found to be present from 2 to 14 years after

resettlement in the country of asylum, depending on the severity and chronicity of the preimmigration trauma endured. The prevalence of PTSD varies among refugee groups involved in different research studies, with rates for refugee groups coming from world regions characterized by unspeakable atrocities being as high as 93%. In addition to PTSD and other mental health problems, some refugees experience survivor guilt for enduring in circumstances where significant others perished. Refugees may also experience various types of physical health problems due to injuries sustained during torture or trauma experiences, interrupted health care access and self-care practices during periods of political unrest or internal displacement in refugee camps, and psychosomatic problems. Accounts of temporary blindness and limb paralysis among severely traumatized refugees exemplify psychosomatic difficulties.

Sojourners

A group of sojourners that has been identified as particularly vulnerable to the development of mental health problems is international humanitarian workers. Their relocation to world regions plagued by natural disasters or political turmoil places them in highly stressful environments. They are faced with multiple, simultaneous adjustments during their relocation experience. First, they are often exposed to situations involving mass trauma, injury, and fatalities, without the level of formal infrastructure support or medical equipment and facilities that are required to effectively respond to the needs of the resident population. Second, they are exposed to a new culture and a new language. Communication barriers may challenge their efforts to provide quality support or mental or physical health services to the resident population under highly constrained environmental conditions. Third, the sojourning process for humanitarian workers often involves separation from their immediate family members and other sources of social support, limiting their coping resources.

One of the major emotional risks of working in traumatic situations is being personally affected by the trauma that others have endured. Changes in personal beliefs about safety and stability in the world, feelings of numbness, memories of troubling events or situations described by others, and intense emotional suffering may occur upon witnessing the suffering of another. Emotional preparation and self-care for humanitarian workers are essential for maintaining

their ability to provide assistance to others in highly stressful circumstances characterized by vicarious exposure to trauma. Follow-up counseling may also be needed after they return to their home countries, bringing with them traumatic memories of their experiences abroad.

Protective Factors

Despite being exposed to multiple, compound stressors across various stages of immigration (predeparture phase, flight or departure phase, relocation phase, and return to country of origin), many immigrants, refugees, and sojourners do not experience serious mental health problems. Members of some migrant groups have even reported an increase in mental or emotional well-being due to being strengthened by adverse life experiences. In recent years, there has been an increased interest in factors that promote mental health besides factors that relate to mental illness. Three key protective factors have become apparent in the lives of migrants who show resilience to multiple life stressors. One of these factors is spirituality or religiosity. Belief in something greater than oneself may assist people in meaning-making about difficult life events and help them to feel some level of control over otherwise uncontrollable circumstances. Another factor is social support. Support may be manifested in the presence of other members of one's immigrant/refugee group in the region of resettlement, immigration agency programs that connect migrants with host families to assist with integration, and so on. Support assists people to cope with and make the adaptations needed to successfully manage the immigration process. An acculturation strategy of integration has also been found to promote positive mental and physical health and to reduce stress levels, when the national policy of the country of settlement permits this strategy. Pursuing integration may allow migrants to access the supports and services of both the host group and their own cultural group. It is ultimately the balance or imbalance of stressors and protective factors in the lives of migrants that shapes their immigration experiences.

Noorfarah Merali

See also Acculturation; Asian Americans; Bilingualism; Cultural Diversity; Hispanic Americans

Further Readings

Al-Issa, I., & Tousignant, M. (Eds.). (1997). *Ethnicity, immigration, and psychopathology.* New York: Plenum.

Berry, J. W. (2001). A psychology of immigration. *Journal of Social Issues*, *57*, 615–631.

Church, A. T. (1982). Sojourner adjustment. *Psychological Bulletin*, *91*, 540–572.

Gaw, K. F. (2000). Reverse culture shock in students returning from overseas. *International Journal of Intercultural Relations*, *24*, 83–104.

Holzman, W. H., & Bornemann, T. H. (Eds.). (1990). *Mental health of immigrants and refugees.* Austin, TX: Hogg Foundation for Mental Health.

Koenig, H. G. (1997). *Is religion good for your health? The effects of religion on physical and mental health.* New York: Haworth.

Kuo, W., & Tsai, Y. (1986). Social networking, hardiness, and immigrant mental health. *Journal of Health and Social Behavior*, *27*, 133–149.

Mollica, R. F., Caspi-Yavin, Y., Lavelle, J., Tor, S., Yang, T., Chan, S., et al. (1996). Manual of the Harvard Trauma Questionnaire. *Torture*, *1*(Suppl.), 22.

Reitz, J. G. (2001). Immigrant success in the knowledge economy: Institutional change and the immigrant experience in Canada, 1970–1995. *Journal of Social Issues*, *57*, 579–614.

Ryan, M. E., & Twibell, R. S. (2000). Concerns, values, stress, coping, health, and educational outcomes of college students who studied abroad. *International Journal of Intercultural Relations*, *24*, 409–435.

Sandoval, J., Frisby, C. L., Geisinger, K. F., Scheuneman, J. D., & Grenier, J. R. (Eds.). (1998). *Test interpretation and diversity: Achieving equity in assessment.* Washington, DC: American Psychological Association.

Silove, D., Sinnerbrink, I., Field, A., Manicavasagar, V., & Steel, Z. (1997). Anxiety, depression and PTSD in asylum-seekers: Association with pre-migration trauma and post-migration stressors. *British Journal of Psychiatry*, *170*, 351–357.

Wagner, R. K., Francis, D. J., & Morris, R. D. (2005). Identifying English language learners with learning disabilities: Key challenges and possible approaches. *Learning Disabilities Research & Practice*, *20*, 6–15.

Wehrly, B. (1995). Ethnic identity development over the life span. In B. Wehrly (Ed.), *Pathways to multicultural counseling competence* (pp. 65–78). Pacific Grove, CA: Brooks/Cole.

INCLUSION

Inclusion refers to educating students with disabilities together with their nondisabled peers. It is an outgrowth of several strands in education (mainstreaming, the Regular Education Initiative) and disabilities studies (normalization), and is often conceived as a civil rights movement comparable to the elimination of segregation by race.

Definitions

Inclusion can be conceptualized as a philosophical approach to education for children with disabilities, or as a specific set of practices to support participation in general education settings. Recently, the focus of inclusion has shifted from access to quality education. Currently, there is no agreed-upon definition of the term. However, inclusion, sometimes referred to as *full inclusion*, incorporates a number of principles that help illustrate what is meant by the term, such as (a) shared responsibility among all school staff for the education of every child, regardless of disability status; (b) all students educated in their neighborhood schools and assigned to classrooms based on their age/grade level; and (c) following the principle of natural proportions, meaning that the proportion of students with disabilities in any given classroom is representative of the community. Additionally, inclusion assumes that all schools are physically accessible and follow developmentally appropriate practices, and that inclusion is implemented across all grade levels and schools. Finally, inclusion generally means that all students should share the same schedule and activities, including co-curricular activities.

Inclusion should not be considered simply the placement of children with special educational needs in general education classrooms, a process known as *dumping*. Neither should it be confused with an earlier iteration known as *mainstreaming*, whereby students with disabilities left a specialized education setting to attend selected classes in the general education setting. Rather, inclusion calls for the development of a universally designed system with the capacity to support all learners and with the infrastructure for the delivery of special education supports and services in the general education classroom.

Current Legal Issues

Legal mandates regarding students with disabilities specify certain rights and entitlements that bear on the issue of inclusion.

The Americans with Disabilities Act (ADA) (1990) requires all public institutions to provide equal

access to disabled and nondisabled persons. The ADA provides a baseline via reasonable accommodations, but does not mandate remediation. Therefore, students who only require accommodations such as ramps or sign language interpreters in order to access educational services are protected by the ADA as well as Section 504 of the Rehabilitation Act of 1973. However, students with disabilities from birth through age 21 have additional rights under the 2007 Individuals with Disabilities Education Improvement Act (IDEIA).

IDEIA, the most recent reauthorization of special education law, does emphasize that consideration of access to the general education curriculum and placement with general education peers should be a major consideration in Individualized Education Plan (IEP) development. However, should attainment of goals and objectives not be achieved satisfactorily in the general education environment with supplementary aids and services, more restrictive placements must be available. The legal definition of least restrictive environment (LRE), therefore, does *not* imply that all students must be educated with their general education peers; rather, it specifies that the student be placed in the least restrictive environment in which educational goals can be achieved satisfactorily, therefore requiring that a full continuum of services be available. Students' IEPs for inclusive services should detail the modifications and supports necessary to meet objectives in the general education setting.

No Child Left Behind (NCLB) includes most students with disabilities in general education assessment for attainment of educational goals. In addition, students with disabilities as a group have their scores disaggregated in order to ensure that they are achieving educational progress at an appropriate level. Students with disabilities are entitled to testing accommodations, which generally fall into the following four categories: (1) how the test material is presented; (2) how the student responds; (3) scheduling and length of test periods; and (4) setting.

Approximately 1%–2% of students with the most severe disabilities are assessed by alternative methods, assuming that these students will have different goals and objectives from the general education outcomes being assessed.

Infrastructures and Supports

Because of the comprehensive nature of the changes necessary, inclusion cannot be implemented as a "special education" reform, but must be implemented as a school- or district-wide reform. In 1992, the National Association of State Boards of Education (NASBE) identified inclusion as the preferred practice model and called for substantive reconfiguration of current practices in order to achieve the creation of a belief system that does not differentiate between disabled and nondisabled students, the merging of special and general education systems, and the abolition of the link between funding special education and the disability label.

Information from a national panel of experts in inclusion used the Delphi technique to identify 92 best practices for successful inclusion. The practices were divided into categories such as Policy; Stakeholder, Leadership, Support, Involvement; Resources and Supports; Professional Practices; Curriculum and the Classroom; and Accountability of Inclusion Programs. Although Policy was deemed the most critical broad area, Resources and Supports was the area with the most practices deemed necessary for a successful inclusion program. Practices seen as necessary included general education and special education teachers (or related service providers) co-teaching classes as equal partners, with sufficient release time weekly to plan together; collaborative teams, including the parents and students, meeting regularly; regular staff development, training, and opportunities to visit successful programs; access to specialists, including speech/language pathologists, occupational and physical therapists, counselors and school psychologists, and reading specialists; and technological and classroom modifications, including augmentative communication devices, adapted computer equipment, specialized furniture, and Braille materials.

In-service training for staff, as well as preservice preparation, should address areas such as the design, implementation, and evaluation of individualized interventions; assessment, including alternative and authentic assessment; working effectively with families; and collaboration.

Successful system change to build the infrastructure required for effective inclusive practices usually involves a system-level, problem-solving approach. Some form of preparatory communication and information dissemination to increase stakeholder buy-in usually precedes implementation of systems-level change. The first phase involves assessing the current levels of system functioning across domains, and the

magnitude of the gap between current status and the goal of inclusion is measured. Next, barriers and supports are analyzed. Plans addressing barriers and supports are developed, often after research on successful models has been explored. Most importantly, a formative and summative evaluation plan is implemented to make adjustments as needed and provide outcome data.

Effective inclusive practices also require a plan of established collaboration with professionals and agencies outside of the school. Health care providers, social service agencies, and the juvenile justice system are examples of professionals who also serve children with disabilities and make decisions and recommendations to families that can affect inclusion.

Research Outcomes

Reviewing the research on inclusion is difficult, as many studies do not define the supports and services provided, but simply focus on where students with disabilities are educated. Nonetheless, some conclusions can be drawn.

Students With Disabilities

Outcomes for students with moderate to severe disabilities in inclusion programs are almost universally strong in both academic and social arenas. When provided with appropriate supports, students have more academic engaged time and do at least as well on standardized tests. They have more opportunities for social interaction, although structured teaching of social skills may be necessary, and show more independence. Some research indicates less emphasis on community and vocational skills, so these may need to be emphasized over time. As students get older, inclusion into the wider community and workforce may be an appropriate addition to inclusion in the classroom.

Outcomes for students with mild disabilities are less clear, perhaps because they are more likely to be placed in general education without significant support. Inclusive programs may raise academic expectations for students with mild disabilities, but students may not reach those expectations without specialized instruction. Some research shows that this can be done successfully in the general education classroom, whereas other research shows greater efficacy with a pull-out model. Socially, students with learning disabilities may be accepted, whereas students with emotional and behavioral disorders are often rejected. Findings on self-concept and self-esteem vary.

Students Without Disabilities

In general, inclusion programs can be beneficial for nondisabled peers, especially when curriculum changes and supports can benefit low achievers. Higher-achieving students may not do as well if the range of expectations in the class is lowered. If too many students with disabilities are included into a single class (which often occurs to allow special services to be targeted to one class per grade level), the effect on nondisabled classmates can be negative, probably because of the lack of attention they receive. In addition, if an included student with disruptive behaviors does not have an effective behavior management plan, this can be harmful to classmates. These findings demonstrate the importance of maintaining natural proportions of students with disabilities and providing sufficient supports and services in inclusive programs. Socially, nondisabled students benefit from exposure to students with disabilities by becoming more accepting and compassionate.

Tiered Model Approach

Current special education law encourages the use of a tiered system of supports, starting with universally designed environments, adaptable curricular materials, and differentiated instructional practices that can be flexibly applied across varying levels of abilities and needs. This supports the underlying principle of educating all learners in their home school in their age/grade classroom. In the second tier, targeted interventions are provided for students requiring more instruction, practice, or adaptations to achieve learning goals. This can include concentrated small group work involving consultation or direct support from student support personnel, such as school psychologists and special educators. The third tier involves increasing intensity and number of supports. Some tiered models consider the third tier to be special education service delivery, where due process is activated. Other models consider special education to begin at a fourth tier. In either case, special education is seen as increasing levels (tiers) of support rather than removal from the general education milieu.

Culturally Responsive Practices in Inclusion

According to NASBE, no single ethnic group will constitute a numerical majority of the school-aged population by the year 2040. Culturally responsive practices involve acquiring knowledge about individuals and their culture, and then developing and implementing procedures and supporting attitudes with the aim of improving services and providing better outcomes. When providing inclusive education, this means that school staff need to understand the attitudes about disabilities, education, and special education within different cultures in the school, and use appropriately trained interpreters and/or cultural brokers to ensure that interventions, goals, and procedures for goal attainment are congruent with the culture of the students.

Catherine A. Fiorello, Jean A. Boyer, and Rebecca J. Thompson

See also Disabilities; Learning Disabilities; Special Education

Further Readings

Kauffman, J. M., & Hallahan, D. P. (Eds.). (1995). *The illusion of full inclusion: A comprehensive critique of a current special education bandwagon.* Austin, TX: Pro-Ed.

McGregor, G., & Vogelsberg, R. T. (1998). *Inclusive schooling practices: Pedagogical and research foundations.* Philadelphia: Allegheny University of the Health Sciences.

National Association of State Boards of Education. (1992). *Winners all: A call for inclusive schools: The report of the NASBE Study Group on Special Education.* Alexandria, VA: Author.

National Association of State Boards of Education. (2002). *A more perfect union: Building an education system that embraces all children.* Alexandria, VA: Author.

Stainback, S., & Stainback, W. (1992). *Curriculum considerations in inclusive classrooms: Facilitating learning for all students.* Baltimore, MD: Paul H. Brookes.

Vantre, E. M. (2001). *Evaluating the presence of best practices for inclusion: The Marsh Inclusion Scale.* Unpublished doctoral dissertation, Temple University.

Williams, B. B. (2006). *Culturally competent mental health services in the schools: Tips for teachers. Helping children at home and school II: Handouts for families and educators.* Bethesda, MD: National Association of School Psychologists.

Zionts, P. (1997). *Inclusion strategies for students with learning and behavior problems: Perspectives, experiences, and best practices.* Austin, TX: Pro-Ed.

INDIVIDUAL DIFFERENCES

Reading involves a complex and ordered set of processes and behaviors that leads to a mental representation that conveys the underlying meaning of a text. Individual differences exist to the extent that readers are proficient in these processes and the extent to which systems are in place to support their occurrence. However, these processes and behaviors are not independent of one another. For example, working memory resources for student readers who are not proficient at lower-level aspects of reading will be consumed by these processes, which will compromise their ability to engage in higher-level processes. As such, it is critical for researchers and educators to identify and diagnose these multiple components of reading in order to help struggling readers.

Being able to identify students at risk of being poor readers is a primary goal of educators and researchers. Conventional wisdom tells us that students must be competent readers to succeed both academically and in society. Despite best efforts to help struggling readers during primary education, the RAND Report, *Reading for Understanding*, provided compelling evidence that there is a crisis regarding the literacy skills of students in the United States. The gaps in performance on measures of literacy competencies between students in the United States and other industrialized countries are widening. At the same time, the gaps between different demographic groups within the United States remain unacceptable. This crisis worsens when students reach secondary education, where texts become more challenging and expectations of gaining and using knowledge from text increase. Currently, there is little emphasis on teaching students to develop higher-level literacy skills that promote deep learning, leaving struggling readers at even greater risk academically.

Developing successful interventions to help struggling readers requires that one have an understanding of what is involved in reading and be able to identify a student's proficiencies in the various aspects of reading. Reading involves a series of cognitive processes and activities that can be conceptualized as

ranging from relatively low level (e.g., phonological decoding) to high level (e.g., generating inferences based on background knowledge). These processes build upon one another, such that the products of one process must be available for the next process to occur. For example, a reader has difficulty accessing the potential meanings of a word unless he or she has accurately accessed the phonological sound of that word. Similarly, a reader cannot determine how a sentence is related to the prior text unless he or she has accurately understood the meaning of that sentence.

The goal of this entry is to discuss factors that influence individual differences in reading comprehension proficiency. Readers struggle for a variety of reasons. As Charles Perfetti argued, readers may struggle because they have not adequately developed proficiencies in lower-level language processes, such as phonological decoding, accessing the appropriate meaning of words, or understanding the grammatical relationships between words in a sentence. Although proficiencies in these lower-level processes are necessary for good comprehension, Jane Oakhill has argued that they do not guarantee it. That is, deep comprehension requires processes and strategies that are conceptually higher order, such as determining how a sentence is related to the prior text or generating appropriate inferences. Finally, students have a number of attributes that support reading comprehension, such as prior knowledge and working memory capacity. These "extratextual" factors can also contribute to individual differences during comprehension. This entry contains a discussion of lower-level reading processes, higher-level processes, postreading processes, and extratextual factors that contribute to individual differences in reading comprehension ability.

Factors That Contribute to Individual Differences in Reading Comprehension

One must first describe what comprehension involves before there can be a discussion of individual differences in reading proficiencies. Comprehension does not equal the process of translating combinations of letters into the sounds that they represent. Many students have had the experience of going through the mechanics of reading a text (even experiencing the sound of the words in their head as they read), but having very little to show for that time in terms of deep comprehension. This anecdote illustrates the distinction between lower-level processes associated with the mechanics of reading and higher-level processes that lead to the construction of a mental representation that is a basis for comprehension.

There are many reasons why a student may struggle to comprehend when reading. Some of these stem from deficiencies associated with processes and skills involved during reading. Psychologists such as Karl Haberlandt have described reading as consisting of levels of processes that build upon one another. These processes can be described as operating on the word, sentence, and discourse levels of analysis and are thought to occur sequentially. Word-level processes make sentence-level processes possible, and sentence-level processes in turn make discourse-level processes possible. It is important to note that there is some debate as to whether these processes are truly sequential or if they occur in parallel. Nonetheless, the end result of these processes working together is the construction of a mental representation of a text that reflects its underlying meaning.

Word-Level Processes

Early stages of literacy development involve fostering the development of several prerequisite skills at the lowest level of linguistic processing, the word level. Readers must be able to decode letters in print, match them with the corresponding sounds they represent, and combine them into syllables that make up words. These skills combine to form a construct called *phonological awareness.* Early success in reading is contingent on readers becoming fluent in these phonological processes. Fluency in decoding is a reader's ability to rapidly and automatically decode whole words and is often measured in words read per minute. A reader is expected to be proficient at these lower-level skills and working on becoming a more fluent reader by the time he or she reaches third grade.

Another aspect of word-level processing involves accessing the appropriate meanings of words. The mental lexicon refers to that part of the mind that stores the word knowledge that makes up a person's vocabulary. After readers accurately identify the phonological symbol that the written word represents, they must access, retrieve, and select the appropriate meaning of the word from the mental lexicon. This issue becomes complicated because words often have multiple meanings. For example, the word *bank* can mean a financial institution or it can mean the land on the edges of a river (of course, there are other

definitions of this word). When reading the sentence, "John took his check to the bank," readers must select the former meaning in order to understand how this word fits into the sentence context. One may assume that the sentence context influences access such that only the appropriate meaning of a word is activated. However, considerable research demonstrates that, initially, all meanings of a word known to a reader are activated, and within about half a second after recognizing a word, the appropriate meaning is selected.

Individual differences in mental lexicon size and lexical access have implications on reading proficiencies. First, as shown by Perfetti and others, the size of a reader's vocabulary knowledge has very important implications for comprehension. If a reader has a small vocabulary, the likelihood of encountering a word not present in the lexicon is high, which compromises a reader's ability to comprehend what he or she is reading. Second, Morton Gernsbacher has also shown that one source of problems faced by struggling readers is difficulty in selecting the appropriate meaning and deactivating the inappropriate meaning. Maintaining inappropriate meanings of words not only consumes limited resources of readers, but also increases the likelihood that an incorrect interpretation will be constructed. This issue is more difficult to remedy because activation and selection are processes outside a reader's conscious awareness.

The National Reading Panel compiled a large body of research that suggests that there are severe ramifications for not developing adequate proficiencies at word-level processes early during literacy education. Several studies have shown that readers who are not proficient at phonological decoding by the end of third grade are likely to have poor reading skills later. Furthermore, Joseph Jenkins and his colleagues have found that when reading words in context out loud, skilled readers read three times more correct words per minute than less skilled readers, and did so with fewer errors. This later finding suggests that readers who are proficient at phonological awareness are able to process texts with less effort, in a shorter amount of time, and with more accuracy than readers who are not proficient at this aspect of reading.

Sentence-Level Processes

Sentence-level processes involve deriving the semantic representation referenced within a sentence.

There are at least two levels of processes that support this endeavor. The first sentence-level process is syntactic parsing, which involves determining the grammatical class of words (i.e., nouns, verbs, articles, prepositions) and establishing their relationships within a sentence. This involves identifying and constructing the syntactic structures that convey noun and verb phrases in a sentence. Syntactic parsing is guided by expectations and experiences within a given language. In English, a subject-verb-object structure is adopted so that one expects the first noun phrase to be the subject, the first verb to be the main verb of the subject, and the object to be the object of the main verb. Violation of this expectation or ambiguities in syntactic structure can result in longer processing and thus increased sentence reading times.

A second sentence-level process involves the underlying meaning of a sentence, which psychologists call the *proposition.* Proposition construction involves grouping contextually appropriate semantic words into meaningful units and reflects the deep semantic relationships between the constituents in the sentence. Propositions consist of a predicate (verb, adverb, prepositional phrase) and one or more arguments or nouns. For example, the sentence "Joe hit the ball" involves the predicate *hit*, which determines the relationship between the nouns *Joe* and *ball.* The propositional structure, along with the activation of the appropriate word meanings, provides the basis for constructing the underlying semantic meaning of a given sentence.

Both word- and sentence-level processes are considered to be lower-level reading processes. Perfetti's verbal efficiency theory assumes that proficiencies in these lower processes greatly influence a reader's ability to comprehend as he or she is initially learning to read or while learning to read in a second language. This is because lower word- and sentence-level processes such as word decoding, lexical access, and proposition construction are consuming the same attentional resources. If lower-level processes are consuming this finite resource space, comprehension will likely suffer. Alternatively, the less effort these processes consume, the more readers can devote attention to aspects of comprehension that are associated with building a memory representation of the textbase and situation model.

Once students have learned to read (primarily after the third grade), individual differences in these lower-level proficiencies manifest themselves most

dramatically when students encounter difficult and challenging texts. This is one reason why students who do not develop adequate proficiencies by the end of traditional literacy education (around the third grade) are at risk of poor performance during later grades. As such, there is considerable emphasis on measuring students' proficiencies in these processes in order to identify at-risk students. Exactly what to measure is an issue of debate, as many school psychologists are now using measures of reading fluency as the primary tools for measuring students' progress during early literacy training, and several have controversially argued that measuring fluency is a proxy for measuring reading skill and comprehension. Alternatively, Perfetti has argued that measuring vocabulary knowledge is correlated with proficiencies at both word- and sentence-level processes.

Discourse-Level Processes

Discourse-level processes support the construction of memory representations that reflect the meaning of a text. This representation is updated upon the comprehension of every sentence. That is, after readers compute the meaning of a sentence, the memory representation for the texts will be updated to accommodate that information. Psychologists such as Walter Kintsch have shown that mental representations of texts contain multiple levels of meaning. Readers construct a representation of the explicit text content, referred to as the textbase. The textbase contains the propositions that reflect ideas that are explicitly described in a text. Comprehension is supported when readers are able to connect these explicit ideas to one another in the mental representation. One way this is done is when the current sentence mentions the same concept referenced in a prior sentence. For example, when a reader encounters a pronoun referring to something already mentioned, he or she needs to understand that these current pronouns refer to arguments previously mentioned in the proposition. In the sentences "Joe hit the ball. He ran to first base." the textbase representation would consist of the propositions [*hit, ball, Joe*] and [*run, first base, he*]. Generating the anaphoric inference that *he* refers to Joe establishes the relationships between these two propositions. Notice, however, that it is up to the reader to infer that the pronoun *he* refers to Joe.

However, establishing that sentences mention the same concepts is most certainly not enough to support

comprehension. Readers may have to rely on their general knowledge of the world to compute inferences that establish the connections between ideas. For example, "John hit a home run. The pitcher was yanked from the game." does not contain the same characters or event. However, most readers would have no trouble inferring a causal relationship between these two events given their knowledge of baseball.

Psychologists believe that coherence emerges with the construction of a situation model, which conveys the underlying situation implied by the explicit text. The situation model is constructed when readers generate inferences that are based on their world knowledge, which enables them to establish implied relationships between text constituents. The situation model establishes relationships between textbase propositions along a number of dimensions, such as agents and objects, temporality, spatiality, causality, and intentionality. It is important to note that both the textbase and situation model representations are part of a highly integrated network that reflects the underlying meaning of a text. As such, it can be difficult to determine the quality of these aspects of comprehension independently of one another. What is important here is that one understand that assessing memory or understanding of the explicit text will not measure a student's comprehension adequately.

There is some controversy as to whether or not discourse-level comprehension arises from effortless, bottom-up processes or through active, constructive processes. Although it is likely that both support comprehension, it may be that the amount of effort and work required to comprehend is contingent on the difficulty of the text or whether or not the reader has a learning goal.

There is a growing body of evidence that skilled and less skilled readers differ in terms of the extent to which they establish relationships between sentences during reading. For example, Joseph Magliano and Keith Millis have explored differences between skilled and less skilled readers by having them produce verbal protocols while reading. Readers are instructed to report their thoughts regarding their understanding of the texts after target sentences. These target sentences are chosen because they have been identified as being causally related to prior text sentences, and as such comprehension of the text is contingent on readers establishing these relationships. They find that skilled readers tend to talk about how a target sentence fits into the prior texts (and in doing so, mention information from the

prior text), whereas less skilled readers tend to talk primarily about the sentence that they just read (and in doing so, mention less information from the prior text). Magliano and Millis argue that these behaviors while producing verbal protocols are indicative of strategies that readers adopt when reading silently. In support of this assumption, Magliano and Millis have shown that the more readers talk about the causally important prior text sentences, the better they tend to comprehend and remember texts that are read silently. Conversely, the more readers talk about the sentence they just read, the poorer they tend to comprehend and remember a text.

Extratextual Factors

Finally, several attributes of the reader support the processes described above and have implications on individual differences in reading skill. One of these is working memory capacity, defined as a cognitive storage and processing system that is of limited size. Marcel Just and Patricia Carpenter have shown that readers with smaller working memory capacities have difficulties dealing with ambiguities at the word and sentence levels. These differences may stem from the fact that working memory is consumed by lower-level language processes for readers with smaller working memory spans. As such, they may not have the resources available to establish the implied relationships between sentences. Additionally, these difficulties stem from the fact that readers with a smaller working memory span cannot maintain multiple interpretations as they are reading the discourse and must choose one, which may not be correct. Readers with higher working memory capacities can maintain multiple interpretations or hypotheses until the correct one is indicated by the discourse. Furthermore, an individual with greater working memory capacity is able to hold the most relevant information in consciousness when completing complex tasks such as reading, resulting in better comprehension of the text.

Another extratextual factor in linguistic processing is metacognition or comprehension monitoring. This is defined as an awareness of understanding or knowing when one needs to be more active when reading. Skilled readers are able to mediate how active their reading level is based upon their perceived level of understanding. Strategies such as self-questioning, summarization of what has been read so far, and mental imagery have all been shown to improve a reader's ability to comprehend. Skilled readers engage in more

self-monitoring and metacognitive thoughts and are therefore more aware of whether or not they understand a text.

Students may differ in their need for cognition, defined as the extent to which individuals enjoy engaging in thinking. As one may expect, students who have a high need for cognition tend to perform better in secondary and higher education. Similarly, Paul van den Broek has postulated that readers differ in their standards of comprehension, which should be related to an individual's need for cognition. Readers who have a high standard of comprehension employ the necessary strategies to acquire a rich, coherent representation of what they read. It may be the case that readers with low standards of comprehension are satisfied if they feel that they understand the individual sentences in a text. However, deep comprehension requires readers to establish how the sentences in a text are related to one another.

Finally, a reader's level of prior knowledge on a topic has implications on comprehension. James Voss and his colleagues have conducted several studies that aptly demonstrate this important point. In one study, they identified students who either had high or low knowledge about baseball. They found that students with high knowledge about baseball demonstrated better comprehension and memory for texts about baseball than readers with low knowledge. However, these benefits in comprehension were not seen in texts that did not pertain to baseball. Additionally, readers with high knowledge of the domain topic, such as biology, are better able to learn from a science text than those readers with low knowledge.

Helping At-Risk Students

At least three approaches are available to address some of the issues described above: skills training, text modification, and increasing exposure to reading. With respect to skills training, one can conceptualize the processes associated with reading as a house of cards. If students lack proficiencies in the earlier stages of processing (the foundation), the entire structure collapses. As one can imagine, remediation of a specific student's problems requires one to accurately identify where problems with reading lie. One problem is that many multiple-choice tests of reading comprehension are not designed to be used as diagnostic tools. Educators and psychologists must be aware of the problems that readers can face during

reading and use assessment tools that tap proficiencies at the different levels of reading processes.

Several interventions are appropriate if a student is not proficient at lower-level reading processes. For example, slow, halting readers have shown improvements in overall reading fluency rate through the use of repeated reading skills practice interventions where single passages are read multiple times, and these results have generalized to other texts. Reading passages multiple times requires students to practice decoding difficult phonetic sequences within a single word and/or multiple words. As another example, Margaret McKeown and Isabel Beck have showed that students who were taught new vocabulary words showed more accurate word knowledge, faster lexical access, and greater overall comprehension than did students with whom they were matched in vocabulary knowledge and comprehension skill prior to vocabulary training. Presumably, these interventions are effective because they lead to a decrease in the working memory resources required to perform lower-level processes, thereby making more resources available for readers to engage in higher-level processes.

However, focusing only on teaching and assessing lower-level processes at the expense of other aspects of reading may be a mistake. Oakhill and colleagues provided data to show that developing proficiencies at this level is necessary but not sufficient for the development of good comprehension. Several interventions are aimed at teaching students to be active readers who focus on developing proficiencies in discourse-level processes. For example, Danielle McNamara has developed an intervention that teaches students to self-explain while reading. Students are taught to ask themselves why events are happening in the context of a text or why an author mentions a particular piece of information. These explanations enable readers to establish connections between sentences and make links to relevant knowledge of the world.

Another approach to assist students at risk is improving the readability of a text. Increasing overlapping terms and explicit relationships in the text benefits low-knowledge and low-skill readers. However, a complexity exists in that readers of differing skill respond to this approach differently. Evidence suggests that low-skill readers benefit from making texts more cohesive and coherent, but skilled readers show no benefit or are hindered by texts that are too cohesive, showing that perhaps skilled readers perform best when challenged by the text.

Finally, increasing a reader's text exposure can affect his or her overall reading skill. Research by Keith Stanovich and colleagues provides evidence that variation in readers' exposure to print account for individual differences in reading performance. When decoding ability was controlled for, variation in exposure to print in both children and adults predicted variation in text decoding. When comprehension ability in third graders was controlled for, print exposure was related to individual differences in fifth-grade comprehension ability.

Joseph P. Magliano and Paul J. Perry

See also Dyslexia; Gifted and Talented Students; Learning Disablities; Learning Strategies; Metacognition and Learning; Motivation; Reading Comprehension Strategies

Further Readings

Alexander, P., Kulikowich, J. M., & Schulze, S. K. (1994). How subject-matter knowledge affects recall and interest. *American Educational Research Journal*, *31*, 313–337.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, *42*, 116–131.

Carver, R. P., & Hoffman, J. V. (1981). The effect of practice through repeated reading on gain in reading ability using a computer-based instructional system. *Reading Research Quarterly*, *16*, 374–390.

Chi, M. T., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, *13*, 145–182.

Cipielewski, J., & Stanovich, K. E. (1992). Predicting growth in reading ability from children's exposure to print. *Journal of Experimental Child Psychology*, *54*, 74–89.

Clifton, C., & Duffy, S. A. (2001). Sentence and text comprehension: Roles of linguistic structure. *Annual Review of Psychology*, *52*, 167–196.

Coté, N., & Goldman, S. R. (1999). Building representations of informational text: Evidence from children's think-aloud protocols. In H. van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 169–193). Mahwah, NJ: Lawrence Erlbaum.

Cunningham, A. E., Stanovich, K. E., & Wilson, M. R. (1990). Cognitive variation in adult college students differing in reading ability. In T. H. Carr & B. A. Levy (Eds.), *Reading and its development: Component skills approaches* (pp. 129–159). San Diego, CA: Academic Press.

Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII* (pp. 559–586). Hillsdale, NJ: Lawrence Erlbaum.

Hacker, D. J., Dunlosky, J., & Graesser, A. C. (1998). *Metacognition in educational theory and practice.* Mahwah, NJ: Lawrence Erlbaum.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*, 122–149.

Magliano, J. P., Millis, K. K., Ozurur, Y., & McNamara, D. S. (in press). A multidimensional framework to evaluate assessment tools. In D. S. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies.* Mahwah, NJ: Lawrence Erlbaum.

Magliano, J. P., Zwaan, R. A., & Graesser, A. C. (1999). The role of situational continuity in narrative understanding. In S. R. Goldman & H. van Oostendorp (Eds.), *The construction of mental representations during reading* (pp. 219–246). Mahwah, NJ: Lawrence Erlbaum.

Myers, J., & O'Brien, E. (1998). Accessing the discourse representation during reading. *Discourse Processes*, *26*(2 & 3), 137–157.

Oakhill, J. V. (1982). Constructive processes in skilled and less-skilled comprehenders. *British Journal of Psychology*, *73*, 13–20.

Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading.* Mahwah, NJ: Lawrence Erlbaum.

Shaywitz, S. E., Escobar, M. D., Shaywitz, B. A., Fletcher, J. M., & Makuch, R. (1992). Evidence that dyslexia may represent the lower tail of a normal distribution of reading ability. *New England Journal of Medicine*, *326*, 145–150.

Spilich, G. J., Vesonder, G. T., Chiesi, H. L., & Voss, J. F. (1979). Text processing of domain related information for individuals with high and low baseball knowledge. *Journal of Verbal Learning and Verbal Behavior*, *18*, 275–290.

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension.* New York: Academic Press.

Voss, J. F., Vesonder, G. T., & Spilich, G. J. (1979). Text generation and recall by high knowledge and low knowledge individuals. *Journal of Verbal Learning and Verbal Behavior*, *6*, 651–667.

Zwaan, R. A., & Brown, C. M. (1996). The influence of language proficiency and comprehension skill on situation-model construction. *Discourse Processes*, *21*, 289–327.

Individualized Education Program

The Individuals with Disabilities Education Act (IDEA) defines *special education* as specially designed instruction to meet the unique needs of students with disabilities and prepare them for further education, employment, and, as needed, independent living. IDEA requires schools to develop an Individualized Education Program (IEP) as a means to ensure that students with disabilities receive a free, appropriate public education. The IEP provides administrators with proof of compliance, teachers with formalized plans, parents with the opportunity to participate in decision making, and students with an appropriate education.

The concept of the IEP process used today emerged in the 1960s as a means to develop an instruction plan to meet the unique need of students with disabilities, which federal legislation mandated in 1974. A multidisciplinary team builds an IEP that matches instruction and services to the unique needs, strengths, and interests of each student with a disability. Increasing parent and student input into the IEP team deliberation creates better understanding, shared ownership, and improved satisfaction with the IEP process. Over time, reauthorization of IDEA modified IEP components, but the basic concepts remained the same. Today, among other changes, the IEP focuses on student involvement and progress in the general education curriculum, and on developing postsecondary goals and strategies for students nearing transition from school to post-high school life.

History

Prior to federal legislation in the 1960s, students with disabilities had no legal right to a public education; therefore, many public schools denied access to students with disabilities. In 1966, Congress first addressed the inability of students with disabilities to obtain a public education. Revisions to the Elementary and Secondary Education Act established a grant program "for the purpose of assisting the States in the initiation, expansion, and improvement of programs and projects … for the education of handicapped children" (Public Law [P.L.] 89–750, 161, 80 Stat. 1204). Congress followed by adding additional demonstration goals to the 1970 Education of the Handicapped Act to stimulate states' development of educational resources and training of personnel for educating students with disabilities.

Ideas from Professors Lloyd Dunn from Peabody College, Evelyn Deno from the University of Minnesota, and James Gallagher from the University of North Carolina at Chapel Hill provided the foundation for today's special education policies. Dunn wanted instructional practices for students with disabilities to

focus on measurement of skills and progress. Deno conceptualized special education as a means to meet the learning needs of children with disabilities. She believed special and general educators need to work cooperatively to maximize learning opportunities for all students. In 1972, in his article "The Special Education Contract for Mildly Handicapped Children," James Gallagher advocated for the creation of a 2-year renewable special education contract between parents and educators that would contain instructional goals: "The contract, composed after a careful educational diagnosis, would commit the special educational personnel to measurable objectives that would be upgraded on a six-month interval." These ideas pioneered the foundation of the IEP included in the 1974 Education for All Handicapped Children Act (P.L. 94–142), which in 1990 Congress renamed the Individuals with Disabilities Education Act (IDEA). The reauthorized 2004 Individuals with Disabilities Education Improvement Act (P.L. 108–446), as with previous reauthorizations, fine-tuned the IEP.

Although federal legislation had the broadest impact on developing the IEP, a series of court cases beginning in the 1970s established the IEP as the foundation for providing special education services. In 1977, seven families in *Mills v. Board of Education of the District of Columbia* sought action from the U.S. Supreme Court to prevent the District of Columbia from excluding their children from public education because of their disabilities. The Court upheld the federal mandate that students with disabilities be given access to an appropriate education, and that an individualized education program be developed for each student.

Components

The initial Act (P.L. 94–142) mandated eight IEP components to help ensure that educational programs meet the needs of students with disabilities:

1. Current level of a student's educational performance,

2. Annual goals,

3. Short-term objectives,

4. Documentation of the special education services to be provided,

5. Time the student will spend in special education and related services,

6. Time the student will spend in regular education,

7. Dates for initiating service and anticipated duration, and

8. Evaluation procedures and schedules for determining mastery of the objectives.

With only a few changes, these components have remained the same across subsequent reauthorizations.

IDEA 2004 requires that the IEP include a description of how the student's disability affects his or her involvement and progress in the general education curriculum. The IEP must contain measurable academic and functional goals, supported by needed accommodations and/or modifications to enable students to become involved and make progress in the general education curriculum, and to participate in district and state assessments. Additionally, the IEPs of students who will take alternative assessments need to include benchmarks or short-term objectives. Transition planning changed in IDEA 2004 to require that before students reach age 16 (or earlier), IEPs need to include measurable postsecondary employment, education, and if needed, independent living goals based upon student interests and strengths determined in part through transition assessment or functional vocational evaluation. The transition IEP will contain a course of study designed to facilitate student attainment of their postschool goals.

IEP Team

A multidisciplinary IEP team develops an IEP to meet the needs of each student. The team consists of parents, at least one general education teacher if the child participates in general education classes, at least one special education teacher, a school administrator or administrative designee, someone who can interpret evaluation results, other individuals who have knowledge or expertise regarding the child or his or her disability, and the student whenever appropriate. When the IEP team plans to develop postsecondary goals and discuss transition services, students must be invited, or, if not in attendance, the team must solicit and then consider the student's interests and preferences. Representatives from public agencies responsible for providing transition services to a particular student must also be invited. Whereas the IEP team develops and monitors the IEP, general and special education teachers develop and implement the daily activities to attain student goals.

IEP Usage

The IEP serves as the fundamental tool to provide students with disabilities an appropriate public education. The appropriateness of an IEP, in part, depends on goals and services matching the unique needs, strengths, and interests of students with disabilities. Stephen Smith, in one of the few available articles discussing the efficacy of the IEP, suggests that students may not receive the individualized instruction indicated in their IEP. This mismatch between what the IEP directs and the actual services students receive may result in formal parental complaints, which due process attempts to resolve through mediation and hearings.

IDEA empowers parents to take an active role in the IEP decision-making process. Students have always been able to attend their IEP meetings, but they seldom did. Now, schools need to invite students of transition age to attend their IEP meetings. Unfortunately, parents and students typically provide minimal input in the development of the IEP. James Martin, Jamie Van Dycke, Barbara Greene, and colleagues directly observed middle and high school IEP meetings and found that special educators directed the process, dominated the discussions, and at the end report the greatest degree of satisfaction with the meeting. They observed students talking 3% of the observed intervals, parents talking 15% of the observed intervals, and special education teachers talking 51% of the time. In a 3-year study of 393 IEP meetings, Martin, Laura Huber Marshall, and Paul Sale found that students knew less of what to do, talked less, understood less of what was said, and felt less positive about the meetings than any other IEP team member. When students receive instruction in how to participate at their IEP meetings, student participation increases and IEP team members report increased satisfaction with the IEP process.

James E. Martin and Vincent J. Harper

See also Disabilities; Inclusion; Learning Disabilities; School Counseling; Special Education

Further Readings

Deno, E. (1970). Special education as developmental capital. *Exceptional Children, 37*, 229–237.

Dunn, L. M. (1968). Special education for the mildly retarded—Is much of it justifiable? *Exceptional Children, 35*, 5–22.

Gallagher, J. J. (1972). The special education contract for mildly handicapped children. *Exceptional Children, 38*, 517–535.

Martin, J. E., Huber Marshall, L., & Sale, P. (2004). A 3-year study of middle, junior high, and high school IEP meetings. *Exceptional Children, 70*, 285–297.

Martin, J. E., Van Dycke, J. L., Christensen, W. R., Greene, B. A., Gardner, J. E., & Lovett, D. L. (2006). Increasing student participation in IEP meetings: Establishing the self-directed IEP as an evidence-based practice. *Exceptional Children, 72*, 299–316.

Martin, J. E., Van Dycke, J. L., Greene, B. A., Gardner, J. E., Christensen, R. W., Woods, L. L., et al. (2006). Direct observation of teacher-directed IEP meetings: Establishing the need for student IEP meeting instruction. *Exceptional Children, 72*, 187–200.

Rossow, L. F., & Stefkovich, J. A. (2005). *Education law: Cases and materials.* Durham, NC: Carolina Academic Press.

Smith, S. (1990). Individualized education programs (IEPs) in special education—From intent to acquienscence. *Exceptional Children, 57*, 6–12.

INDIVIDUALS WITH DISABILITIES EDUCATION ACT

The Individuals with Disabilities Education Act (IDEA), which is a supersession of the Education of All Handicapped Children Act of 1975 (Public Law [P.L.] 94–142), requires states and their school districts to provide individuals with disabilities a free and appropriate education. This law dictates how states and school districts provide special education and related services to more than 6 million children with disabilities. This entry details the major components of IDEA, a history of IDEA, and a description of key provisions of this legislation.

Major Components of IDEA

The IDEA comprises four parts: A, B, C, and D. Part A of IDEA, General Provisions, details the purpose of the law. Part B, Assistance for Education for All Children with Disabilities, lays out the requirements for a free and public education in addition to providing federal money to states through IDEA's formula grants. Part C, Infants and Toddlers with Disabilities, requires states to provide services to preschool children with disabilities, while also providing grant money to states to provide services to infants and

toddlers. Finally, Part D, National Activities to Improve Education of Children with Disabilities, primarily functions as a grant-funding entity because it provides financial support for special education personnel preparation, technical assistance, and funding for research projects.

It should be noted that although this law does provide some funding, it does not come close to completely funding the costs that states incur for providing services to students with disabilities. Although P.L. 94–142 stated that Congress would fund IDEA at a rate up to 40% of the excess costs for educating children with disabilities, the federal government has never reached the 20% level.

Education for All Handicapped Children Act of 1975

Until the passage of the Education for All Handicapped Children Act of 1975 (P.L. 94–142), the educational needs of children with disabilities were not being met. This was the first law that required a free and appropriate education to be provided to all children, regardless of their disability. Unlike previous statutes, P.L. 94–142 was a funded mandate that provided greater incentive to states to obey this law. As a direct result, states were required to provide services for students between the ages of 3 and 21. Although the enactment of this law was a step in the right direction, there were still barriers to educating all children. As a result, this law has been amended many times.

Amendments of 1986

In 1986, P.L. 99–457 was passed, which amended P.L. 94–142. This amendment required states to provide a free and appropriate education to children with disabilities from birth to 2 years old. In addition, this amendment provided financial incentives for states to offer additional services for preschool children. As such, the federal government added Part H to the law, which established grant mechanisms for servicing infants and toddlers with disabilities.

Individuals with Disabilities Education Act of 1990

In 1990, Congress reauthorized P.L. 94–142 and renamed it the Individuals with Disabilities Education Act (P.L. 101–476), also known as IDEA. Under this revision, key provisions to services were added. Specifically, the word *handicap* was replaced with *disability*, signifying that a disability is a natural part of the human experience. In addition, it mandated transition planning to help students make the transition from high school to community participation, postsecondary education, and the workforce. This reauthorization also required Individual Education Program (IEP) teams to consider assistive technology devices when creating a student IEP. In regard to children from birth to 2 years of age, Part H was changed to Part C. This allowed states to establish and implement programs for early intervention services for children at risk for disabilities.

Individuals with Disabilities Education Act of 1997

Although the majority of tenets established under P.L. 94–142 were maintained under the 1997 reauthorization, there were significant alterations. Under the reauthorized federal special education law (P.L. 105–17), special education officials were required to increase their efforts to educate children in the least restrictive environment possible. Specifically, educators must document the extent to which a student will participate in the general education curriculum with his or her nondisabled peers. Second, this reauthorization added related services (i.e., counseling, physical therapy, occupational therapy) to the types of services to be provided for transition activities. In other words, services should be based on the individual student's needs in addition to preparing him or her for employment and independent living.

Individuals with Disabilities Improvement Education Act

In 2004, IDEA was again reauthorized. As with previous P.L. 94–142 reauthorizations, there were several modifications to the law. One of the major and highly debated modifications was the change in learning disability eligibility criteria. In the latest rendition of IDEA, local education agencies are not required to take into consideration whether a child has a severe discrepancy between achievement and intellectual ability when determining if a child has a specific learning disability. Better known as the discrepancy

model, this form of identification has been criticized because it does not lead to interventions and the fact that it relies too heavily on IQ test scores. In its place, the local education agency may use a process that determines if the child responds to scientific, research-based intervention as a part of the evaluation procedures. This process, also known as Response to Intervention, is not without its critics as well. As a result, the process for identifying learning disabilities will continue to be highly debated.

In addition to modifying identification procedures, IDEA was made to closely align to another federal education law, No Child Left Behind (NCLB). In accord with NCLB, special education teachers must now be highly qualified. In this instance, "highly qualified" means that all special education teachers must have a state special education certificate and may not hold an emergency, temporary, or provisional certification. In addition, if these teachers are responsible for content courses, they must be licensed in the subjects taught.

Eligibility for Services Under IDEA

To be eligible for services under IDEA, a child must have a specific disability that has an adverse effect on the student's performance. The following disabilities are recognized by IDEA:

Autism

Deaf-blindness

Deafness

Emotional disturbance

Hearing impairment

Mental retardation

Orthopedic impairment

Other health impairment (e.g., attention deficit hyperactivity disorder)

Specific learning disability

Speech or language impairment

Traumatic brain injury

Visual impairment

Evaluation for IDEA Services

An evaluation for services under IDEA consists of procedures to determine whether a child has a disability. During this procedure, the child must be assessed in all areas related to the suspected disability. In general, a comprehensive assessment consists of the following:

1. Individual psychological examination (e.g., general intelligence, academic functioning, social emotional functioning, and learning strengths and weaknesses),

2. Social history based on parent and teacher interviews,

3. Academic history based on teacher and parent interviews,

4. Classroom observation of the student,

5. Functional behavioral assessment to determine the relationship between the child's performance and the variables that are related to its occurrence,

6. Other assessments as necessary depending on the specific disability evaluation category (i.e., speech language evaluation, occupational therapy evaluation, bilingual assessment, auditory and visual assessment).

Parental Consent

Before any assessment or student placement, parents must be informed and consent to the evaluation. This is done to protect the legal rights of parents and their children. Parental consent consists of the following:

1. The parent has been informed of all information relevant to the activity for which consent is sought.

2. The parent agrees in writing to the proposed assessment procedures.

3. The parent is notified that consent is voluntary and can be revoked at any period during the assessment procedure.

Individualized Education Program

The Individualized Education Program (IEP) is a written document that details the educational objectives and expectations for students with disabilities. In accordance with IDEA, specific goals must be developed based on the strengths of the child, the concerns of the parents, and the results of the most recent evaluation. Mandatory information in an IEP includes the following:

1. A statement of the child's present levels of academic achievement and functional performance;

2. A statement of measurable academic and functional annual goals;

3. A description of how the child's progress toward meeting the annual goals will be measured;

4. Periodic reports on the progress the child is making toward meeting the annual goals;

5. A statement of the special education, related services, and supplementary aids with which the student will be provided;

6. An explanation of the extent, if any, to which the child will not participate with children without disabilities in the general education classroom;

7. A statement of any accommodations that are necessary to measure the academic achievement and functional performance of the child on state and district-wide assessments;

8. If the IEP Team determines that the child must take an alternate assessment instead of a particular regular state or district-wide assessment of student achievement, a statement of the reasons for that determination; and

9. The projected date for the beginning of the services, in addition to the anticipated frequency, location, and duration of these special education services.

Individual Education Program Team

The Individual Education Program (IEP) team is a multidisciplinary team that is concerned with determining if a child has a disability, and if so, the services that are necessary to ensure a free and appropriate education. This team is mandated by IDEA so that the child is guaranteed that any evaluation is comprehensive and conducted by different professionals to decrease subjectivity. In accordance with federal mandates, the team must use a variety of assessment measures and gather information from multiple sources. Members of the multidisciplinary team generally include the following: parents of a child with a disability; general education teacher; special education teacher; school psychologist; school social worker; guidance counselor; school nurse; physician; parent advocate; and whenever appropriate, the child with the disability.

Least Restrictive Environment

As a result of the passage in 1975 of P.L. 94–142, all children with disabilities were guaranteed the right to be educated in the least restrictive environment. The least restrictive environment refers to the fact that children with disabilities should be placed in the environment that is best suited for their educational needs. By law, students with disabilities should be educated with children without disabilities to the greatest extent possible.

Transition Services

Transition services are a coordinated set of activities designed to provide students with the practical skills and knowledge that will allow them to transition to adulthood. Beginning when the child turns 16, an Individual Transitional Education Plan (ITEP) must be in place for students with disabilities. The ITEP must include the following:

1. A statement of measurable postsecondary goals related to training, education, employment, and independent living skills;

2. A statement of long-term outcomes for the student; and

3. A list of participants involved in the planning and development of the individualized transitional education program.

Importance

Before its inception, millions of children with disabilities were denied a free and appropriate education, and as a result, many students were unable to reach their full potential. From educating all children regardless of their disability to providing transition services to adulthood and changing how society views the capabilities of individuals with disabilities, this legislation has changed the educational landscape.

Scott L. Graves, Jr.

See also Disabilities; Individualized Education Program; No Child Left Behind; Special Education; Standardized Tests

Further Readings

Batshaw, M. (Ed.). (2002). *Children with disabilities* (5th ed.). Washington, DC: Brookes.

Council for Exceptional Children. (2004). *The new IDEA*. Arlington, VA: Author.

Education for all Handicapped Children Act of 1975, 20 U.S.C. 1400 et seq.

Individuals with Disabilities Education Improvement Act of 2004, P.L. 108–446, 20 U.S.C. 1400 et seq.

President's Commission on Excellence in Special Education. (2002). *A new era: Revitalizing special education for children and their families.* Washington, DC: Author.

Smith, T. (2005). IDEA 2004: Another round in the reauthorization process. *Remedial and Special Education*, *26*, 314–319.

Telzrow, C., & Tankersley, M. (2000). *IDEA amendments of 1997: Practice guidelines for school-based teams.* Bethesda, MD: National Association of School Psychologists.

INDUCTIVE REASONING

Inductive reasoning, or induction, refers to inferences from evidence that are more or less plausible; a good inductive inference is likely to be true. In contrast, the conclusions of deductive inferences are guaranteed to be true by the truth of the premises upon which they are based. Deductive inferences occur in mathematical or logical contexts; almost all other judgments involve induction.

Consider a gardener wishing to buy fencing for a square plot 18 feet on a side. Concluding that the perimeter of the plot equals 72 feet is deductive. What it means to be 18 feet on a side is to have a 72 feet perimeter. This deductive inference is an important piece of solving the problem. Ultimately, however, the gardener has to make an inductive judgment about how much fence he or she will need. Constructing a fence always involves some amount of waste and obstructions to be fenced around. In planning a fencing purchase, it is wise not to take the premises as conclusively true; the plot is unlikely to be exactly 18 feet, nor perfectly square. In a mathematics classroom, the problems demand deduction; outside, in the garden, the hard problems demand induction.

Because almost every judgment has some element of induction, the study of induction is a rich and varied enterprise. Questions about the nature and justification of inductive judgments have a long history in philosophy. In psychology and education, there are two major approaches: a focus on form and a focus on content. These perspectives have led to different pictures of the development of induction and different approaches to teaching good inductive practices.

From a formal perspective, induction involves assessment of probability or association. A weather forecaster's statement that "there is a 70% chance of rain tomorrow" reflects the association between current conditions and future rain: On 70% of the days like today, it has rained tomorrow. Statistics and probability theory provide formal means for calculating association and probability (P). Bayes's theorem defines the probability that a hypothesis is true given some data (e.g., that it will rain tomorrow given today's conditions) in terms of the "prior probabilities" of the hypothesis and the data (how likely rain is on any day, how common are today's conditions) and the "likelihood" of the hypothesis—the probability of the data when the hypothesis is true (how often we see today's conditions when it actually does rain the next day).

$$P(\text{hypothesis} \mid \text{data}) = P(\text{hypothesis})$$
$$* P(\text{data} \mid \text{hypothesis})/P(\text{data}).$$

Probability judgments are a basic element of psychology. Even traditional learning theory (behaviorism) assumes that learners can calculate associations. Animals and infants are able to learn quite complex patterns of probabilities. Whether people's intuitive judgments conform to or violate axioms of probability theory is a major focus of debate. One influential theory is that people reason with heuristics that are efficient and effective given the environments typically encountered. The cognitive abilities underlying valid statistical inference may require particular environmental support (e.g., schooling) or developmental achievements (e.g., Piaget's formal operations). Some psychologists have proposed that people should be explicitly taught good principles of statistical reasoning.

In education, a formal approach to inductive inferences has often focused on teaching the scientific method, such as control of variables in experimentation. More generally, critical thinking is a set of skills for evaluating arguments based on formal principles of logic. Students who learn to construct valid experiments, recognize confounded evidence, uncover assumptions, and avoid contradictions will make better inductive inferences.

Formal principles are independent of the particular content (weather, gardens) that is the subject of judgment. Inductive arguments are often characterized as empirical; they derive not from first principles, but from knowledge of particular facts. People with better facts make better judgments. No matter how skilled at critical thinking or probability analyses, a novice will often make worse predictions than an expert. A classic finding in psychology is that skilled reasoners in one domain are often terrible on analogous problems with unfamiliar content. People seem to reason with

informal heuristics built up from experience, or they have causal knowledge and explanatory theories that warrant conclusions. Educators often observe that learning good inferential skills in one domain does not transfer to another.

One way to improve inductive reasoning is to teach content knowledge rather than (or in addition to) general decision-making strategies. An emphasis on core knowledge holds that sound judgments result from having the requisite background knowledge to identify plausible or workable solutions. Advocates of a domain-specific approach to cognitive development suggest that adult performance results not (just) from general improvements in critical thinking or intellectual abilities, but from greater knowledge and better theories. Children are novices in many areas. Making good decisions involves the expertise that is often characterized as common sense.

A content-focused approach to induction emphasizes abduction, or "inference to the best explanation." Content knowledge is required to identify good explanations. Faced with brown, wilted tomato plants, an expert gardener may draw on knowledge about common pests and properties of tomatoes to identify the problem. Although her reasoning may involve an (implicit) application of Bayes's theorem, the appropriate weighting of probabilities (the priors) would be the discriminating feature of expert judgment. Abductive inference involves testing hypotheses against theory ("does this make sense?") as much as testing against data.

Content-focused accounts of inductive inference face the question of how expert knowledge is acquired. One influential suggestion is that much of our inductive success is the result of innate knowledge. Evolution may have equipped us with the priors we need. An alternative is that content knowledge is the result of formal inference principles. Attending to patterns of covariation and probabilities in the environment is a general way to develop specific knowledge. Bayesian networks are one approach to the question of how domain-general principles can yield domain-specific theories. These models have been very influential in machine learning. Whether human learning can be understood along the same lines is an open research question.

Charles Kalish

See also Correlation; Deductive Reasoning; Scientific Method

Further Readings

Gilovich, T., Griffin, D., & Kahneman, D. (2002). *Heuristics and biases: The psychology of intuitive judgment.* New York: Cambridge University Press.

Holland, J. H., Holyoak, K. J., Nisbett, R. E., & Thagard, P. R. (1986). *Induction: Processes of inference, learning, and discovery.* Cambridge: MIT Press.

Koslowski, B. (1996). *Theory and evidence: The development of scientific reasoning.* Cambridge: MIT Press.

INFERENTIAL STATISTICS

Two basic types of reasoning underlie all of science—deductive reasoning and inductive or inferential reasoning. Deduction involves reasoning from a set of starting assumptions or general principles to a specific future observation. Induction requires reasoning from a limited set of observations to draw conclusions about things that have not been observed or to derive general principles. That is, induction (inference) is the process by which researchers attempt to learn what the general principles or scientific laws are.

Deduction will fail if one or more of the general assumptions is false. If one has the following starting assumptions:

All cows give milk.

Ferdinand is a cow.

The deduction that "Ferdinand gives milk" follows logically. If one were to observe Ferdinand and find out that he does not give milk, one of the two starting assumptions must be false. In this case, the most likely reason for the failure of the prediction is that Ferdinand is not a cow but a bull (but note that the first assumption might also not be true).

Science uses deduction to make predictions about future observations based on past observations. Suppose a researcher has a theory (general assumption) that verbal praise will increase the reading level of second graders. The researcher uses deduction to make a prediction about what will happen if verbal praise is used as part of the program of reading instruction for a sample of second graders. The researcher predicts that their reading scores will rise more than the scores of a control group that does not receive verbal praise. That is, the researcher uses

deduction to reason from a theory or set of beliefs to a hypothesis or prediction about future observations.

Now the researcher wants to test the hypothesis that students who receive verbal praise will show more growth in reading scores than students who do not receive praise. Because the researcher cannot administer the treatment to all second graders, he or she must reason from a limited set of observations to a general conclusion about the effectiveness of praise on learning to read. Here, the researcher must use inferential or inductive reasoning to go from a specific set of observations (researcher's data) to derive or support a general principle. Unfortunately, for reasons discussed shortly, if the researcher observes higher reading scores in a group of students who have received verbal praise, he or she cannot proceed directly to the conclusion that praise is effective.

Inferential statistics, as a branch of applied mathematics, provides a way to reason inferentially with quantitative information such as reading test scores. But before discussing inferential statistics in detail, it is necessary to distinguish between two contexts in which one might want to use quantitative data. The basic issue is whether one has data on all of the entities with which he or she is concerned. If, for example, the Board of Education of the state of Texas wants to know the level of mathematics achievement of eighth-grade students in the state in the current year, it would be theoretically possible to test the mathematics achievement of every eighth-grade student and compute the mean, standard deviation, and other statistics. These statistics describe the group that was tested and, disregarding the possibility of measurement or computational errors, the results accurately reflect the mathematics achievement of Texas eighth-grade students. The mean and other statistics computed on the group that has been tested are called *descriptive statistics*. If the Board has no interest in going beyond this specific group, then no inference is involved. However, it may not be practical to test every student for financial or other reasons. In this case, the entire group about which the Board wants to make a statement is called a *population* and the Board could select a subgroup from this population, called a *sample*, for testing and compute the mean, standard deviation, and so on of the sample. The results computed on the sample would be descriptive statistics for the sample, but the Board's objective is to reason or generalize from the sample to the population. Therefore, because they want to use the results from the sample to draw conclusions about the population, the Board must use *inferential statistics.* Whenever one uses the results from a sample to draw conclusions about a larger population, he or she is engaged in statistical inference. Notice that in statistical inference, descriptive statistics from the sample are used to reach conclusions about the population.

The heart of statistical inference is probability. Because the fact that using inferential statistics implies that not all of the cases of interest are measured, there is always the chance that the conclusions drawn from the sample will be incorrect. Researchers would like to know and, if possible, control the probability that a mistake will be made or an incorrect conclusion will be reached.

There are several ways that researchers can ask questions of the data they have gathered or plan to gather. Exactly how they use inferential statistics will be determined by how they phrase the questions. Some ways may be more useful than others for some purposes.

Testing Null Hypotheses

The origin of inferential statistics is bound up with what is called testing a null hypothesis. During research, researchers derive what they expect to find using deduction from earlier work, but because they wish to generalize to unobserved cases, they must use inference. If an individual draws three marbles from a bag containing 100 marbles and all of the marbles drawn are green, it does not necessarily follow that all of the marbles in the bag are green, but if three green marbles are drawn and then a blue one, the individual knows for certain that *not all* of the marbles in the bag are green. This is the logic behind testing null hypotheses. A *research hypothesis* states what the researcher believes is true but cannot assert its truth based on a limited sample. What researchers do instead is develop a hypothesis, called the *null hypothesis*, which, if they can reject it as untrue, implies that the research hypothesis is true. For example, if a researcher believes that children who receive verbal praise read at a higher level than students taught another way (the research hypothesis), then the null hypothesis would be that students taught using verbal praise *do not* read at a higher level. If the researcher conducts a study and obtains data that are unlikely if the null hypothesis is true, then he or she concludes that the null hypothesis is false, and this

gives support for the research hypothesis. That is, if, based on the data, the researcher finds that students taught with praise read at a higher level, he or she has the statement "students taught using verbal praise do not not read at a higher level." The data are incompatible with the null hypothesis, producing the second *not*, and the double negative affirms the positive.

What is done statistically when stating a null hypothesis is to specify the distribution of outcomes one would expect to observe in a very large number of repetitions of an experiment *if the null hypothesis is true*. These outcomes will have different probabilities of occurring. For example, suppose a couple is planning a family of 10 children. There are 11 different numbers of boys and girls that they could have, but these 11 outcomes have different probabilities. If one assumes (the null hypothesis) that the chance of any child being a boy is .5 and the chance of a girl is also .5, the chance (probability) that all 10 children will be boys (or all girls) is about 1 in 1,000, whereas the chance that there will be 5 boys and 5 girls is about 1 in 4. Suppose that after 20 years of marriage, the couple has produced 10 boys. One of two conclusions can be reached: either that the result happened by chance (unlikely with a probability of 1/1,000), or that the probability of a boy at each birth is higher than that of a girl. The null hypothesis was that the probabilities of boys and girls were equal. The outcome observed was *unlikely to have occurred by chance under the conditions specified in the null hypothesis*, so the conclusion that the null hypothesis is not true is reached.

A researcher can reject the null hypothesis if his or her observation is either larger than would be expected or smaller than would be expected. If the null hypothesis is concerned only with differences in a particular direction (for example, a researcher might be concerned only with whether verbal praise raised reading achievement, not with whether it reduced achievement), the researcher rejects the null hypothesis only if the difference was in the predicted direction. This is called a *one-tailed null hypothesis.* If the researcher does not care about the direction, he or she would reject the null hypothesis when he or she observes an unlikely result in either direction, in which case the null is called a *two-tailed hypothesis.*

Testing null hypotheses has been widely criticized and vigorously debated for more than 50 years. The criticisms have ranged from the assertion that the null hypothesis is "quasi-always" false, and therefore

rejecting it gives researchers no information (see Meehl, 1978) to the observation that the question asked by null hypothesis testing is not the question researchers want to ask (Cohen, 1994; Schmidt, 1996). Because null hypothesis testing has had such a central place in statistical inference and continues to be the most widely used inferential procedure, this entry considers the processes involved in some detail. Later, this entry looks at some of the alternatives that have been proposed.

Sampling Distributions and Test Statistics

In a research study such as the one concerning the effect of verbal praise on learning to read, if the researcher expects verbal praise to be effective, the null hypothesis would be that praise will have no effect or a negative effect. This means that if the null hypothesis is correct, the researcher would expect, in a large number of repetitions of a study of praise, that the mean reading score of praised students would be the same as (or less than) the mean of students not receiving praise. In any single study, the means might differ by a small amount. The question that statistical inference procedures attempt to answer in the case of null hypothesis testing is whether the difference the researcher observes is too large to have occurred by chance if the null hypothesis is true. If the probability of getting a result such as the one obtained is less than 5% (or 1%, the conventional critical probabilities), the researcher concludes that the result is statistically significant, that he or she can reject the null hypothesis, and that the results support the research hypothesis. Note that what the phrase *statistically significant* really means is "unlikely to have occurred by chance under the conditions specified in the null hypothesis."

Statistical inference is based on the concept of the *sampling distribution*. Let us conduct a thought experiment. Suppose that the Texas Board of Education wishes to know what the level of mathematics knowledge is among eighth graders. The mean for all students in the state is a population parameter, a number that characterizes the entire population. It is conventional to use Greek symbols to represent parameters, so the symbol μ (mu, the Greek lower-case letter for M) is usually given to the population mean. If the Board measures the mathematics knowledge of a sample of N (for example, 25) students from this population, the result is a descriptive statistic for the sample, but it can be used to draw an inference about the nature of the

population. Roman letters are usually used for sample statistics, so M is the symbol often used for the mean.

In this thought experiment, the Board draws not one sample but a very large number of samples (all of the same size) and computes the mean (and other descriptive statistics) for each sample. They now have as data not a distribution scores for individuals but a distribution of means for samples, each mean based on the scores of N students. This distribution of means is called the *sampling distribution of the mean*, and it is used as the basis for statistical inference about the population mean. The sampling distribution of the mean contains means (Ms) from samples of size N. The other descriptive statistics that might be computed (such as the standard deviation) also have sampling distributions, and inferences about these statistics would be based on those sampling distributions. Because most of the practice of statistical inference is focused on means, the mean will be used for the development of inferential procedures, but readers should be aware that exactly the same logical processes can be applied to standard deviations, correlation coefficients, and other descriptive statistics. The only difference is in the test statistic that is used. In every case, the null hypothesis specifies a value for the parameter of interest and defines the sampling distribution of the statistic, and statistical inference allows researchers to state how likely the particular sample statistical value is if the null hypothesis is true.

Two additional pieces of information are needed to complete this inferential picture. First, we need to know the mean and standard deviation of the sampling distribution under the null hypothesis. Next, we need a way to determine the probability of any particular outcome given the sampling distribution.

Statistical theory has shown that the mean of the means from a large number of independent random samples from a population is equal to the mean of the population. That is, the mean of the sampling distribution, the mean of Ms, is equal to μ. It can also be shown that if the population standard deviation (called σ, Greek lower-case letter sigma for S) is known, then the standard deviation of the sampling distribution of M is

$$\sigma_M = \frac{\sigma}{\sqrt{N}}.$$

When σ is not known, σ_M must be estimated from the data. One way to get a good estimate of σ_M is with the formula

$$\hat{\sigma}_M = \frac{S}{\sqrt{N-1}},$$

where S is the standard deviation computed from the sample and $\hat{\sigma}$ indicates an estimate rather than an exact value.

Now the mean and standard deviation of the sampling distribution are known if the null hypothesis is true. All that remains is to develop a way to determine the probability of any sample result from this sampling distribution. This involves the use of what is called a *test statistic*. For the current example, one test statistic will be needed if σ_M is known, and a different one if $\hat{\sigma}_M$ must be used.

The test statistic when σ_M is known is called Z. The value of Z reflects how far a given observation is from the mean of the distribution in units of the standard deviation of that distribution. It can be computed as

$$Z = \frac{M - \mu}{\sigma_M}.$$

Z will be negative if the observed mean is below μ and positive if it is above μ. Statistical theory tells researchers that Z follows the normal distribution, also often called the *bell-shaped distribution* or *bell curve*. Most statistics books contain a table that includes the probability of obtaining a value of Z of a particular size. For example, the probability of obtaining a Z of $+1.96$ or greater is 2.5%. Therefore, if a researcher conducts a study using a one-tailed hypothesis and obtains a Z for the sample mean of $+1.96$, the researcher can be confident that if he or she rejects the null hypothesis he or she will make an error no more than about 3% of the time because a result this different from the population mean would occur by chance only about two to three times in 100 repetitions. (For a two-tailed hypothesis, the probability from each end of the distribution is included, so the probability would be 5%.) For example, if the null hypothesis says that the mean reading score in the population is 50 or less (assume it is known that the standard deviation in the population is 10) and a mean of 55 in a sample of 25 students is obtained, the value of σ_M is $10/\sqrt{25} = 2$, so the Z statistic for these data is

$$Z = \frac{55 - 50}{2} = +2.5.$$

Consulting the probabilities for the normal distribution, the researcher would find a value of .006. This

means that, on the basis of the results, the researcher can reject the null hypothesis that $\mu \leq 50$ with a probability that he or she is making a mistake of less than .01. A sample mean of 55 is so unlikely to occur in a sample of 25 from a population with a mean of 50 or less that the researcher concludes that the mean is not 50 but some other (unknown) value. Note that rejecting the null hypothesis does not tell what the population parameter is, only what it is not (μ is not less than or equal to 50). This feature is the source of one of the main criticisms of null hypothesis testing.

When σ is not known, the logical process is identical, but the details are a little more complicated. In this case, the test statistic is known as t, but a different probability distribution must be used depending on sample size. That is, for the Z test statistic, there was only one table of probability values, but for t, there is a different table for each possible sample size. The value for t is computed using the formula

$$t_{df} = \frac{M - \mu}{\hat{\sigma}_M}.$$

The df subscript tells the researcher he or she needs to look in a table of probability based on the sample size, df equals $N - 1$, and there is a different probability table for each value of df (most statistics books include a single table of critical values of the t statistic—values with a probability of .10, .05, .025, and .01 for different values of df—because these are the most commonly used values to make a decision about the null hypothesis).

Using the above example, if the researcher did not know the population standard deviation was 10 but had computed that value from the data, $\hat{\sigma}_M$ would be $10/\sqrt{24} = 2.04$ and

$$t_{24} = \frac{55 - 50}{2.04} = 2.45.$$

Using the appropriate table, the researcher would find that the probability of a t this large is just over .01. That is, the researcher would once again reject the null hypothesis, but the risk that he or she is making an error is slightly larger. When the researcher does not know σ, the distribution of the test statistic is slightly wider, so a particular difference from the null value is slightly more probable. The larger the sample size, the smaller the difference between Z and t.

Most research situations involve comparing the means (or other statistics) from two or more samples. In such cases, the null hypothesis almost always is that the samples come from a common population or from populations that have equal parameters. That is, the null hypothesis is of the general form

$$\mu_A = \mu_B = \mu_C \ldots$$

The details of how hypotheses like this are tested can be found in the texts listed in the Further Readings, but the underlying logic remains exactly as outlined. An appropriate test statistic is computed and the researcher determines how likely it is that he or she would get a value that large if the null hypothesis is true. If the value of the test statistic is too large, the researcher concludes that the null hypothesis is false.

Alternatives to Testing the Null Hypothesis

Confidence Intervals

Several alternatives to null hypothesis testing have been proposed as ways to draw valid inferences from research data. The most widely used is called the confidence interval approach. The essence of confidence intervals is that they approach the issue of inference in a positive way, seeking what is true, rather than in a negative way, rejecting what is shown to be false.

Continuing with the thought experiment in which a large number of samples from a population were drawn and the mean for each sample was computed, when a null hypothesis with a 5% chance of making a mistake in rejecting it was tested, a Z score or a t score for the sample result in the distribution specified by the null hypothesis was computed. Alternatively, *critical values* (*CVs*) for the mean in the null distribution could have been computed by using the value of Z or t that includes 95% of the normal or t distribution. The formula for the normal distribution is

$$CV_{NH} = (\pm Z_{.5\alpha})(\sigma_M) + \mu_{NH},$$

where

CV_{NH} is a critical value for rejecting the null hypothesis,

$\pm Z_{.5\alpha}$ is the lower ($-$) or upper ($+$) Z for an interval that has the specified probability α in the normal distribution (note that half of alpha is in each end of the distribution),

σ_M is the standard error of the mean,

μ_{NH} is the mean of the null hypothesis distribution.

An identical formula with appropriate changes for the t distribution can be used when the population standard deviation is not known, and comparable formulas exist for pairs of means and for other statistics such as the correlation coefficient. The result is an interval that includes the specified proportion of the null distribution. If the sample result, M, falls within that interval (called the *interval of nonrejection*), researchers would not reject the null hypothesis, but if M falls outside that interval (called the *rejection region*), researchers would reject the null.

Now suppose that instead of placing the probability interval around the null hypothesis value, researchers place it around the sample result, M. The sample has been drawn from a population that has a mean μ_{Pop}. Means of samples drawn from this population will form a sampling distribution that is a normal distribution with a mean equal to μ_{Pop} and a standard deviation of σ_M. When researchers test a null hypothesis, they are asking whether it is reasonable to believe that $\mu_{Pop} = \mu_{NH}$. If M falls in the interval of nonrejection, then the probability interval that researchers compute around M will include μ_{NH}. However, if the sample mean falls outside the interval of nonrejection, then the interval around M will not include μ_{NH}. The formula is

$$CV_{CI} = (\pm Z_{.5\alpha})(\sigma_M) + M,$$

and the resulting interval is called a confidence interval. Comparable formulas are available when σ_M is not known and for statistics other than the mean.

The essence of confidence interval thinking is that researchers now have an interval that has a specified probability of including the unknown parameter μ_{Pop}. If a researcher draws a very large number of samples from the population and computes a confidence interval around each sample statistic (e.g., M), 95% (or whatever probability the researcher has specified by his or her choice of values for Z) of those confidence intervals will include the population parameter. Thus, whereas with null hypothesis testing, researchers rule out the region of nonrejection for μ_{Pop} when they reject the null, but leave the rest of the possible values as a region of uncertainty, with confidence intervals, the researchers rule out all of the possible values except those within the confidence interval. In this sense,

confidence intervals give researchers information about where the parameter is, rather than where it is not.

Another advantage claimed for confidence intervals is that they subsume the null hypothesis test. That is, if the confidence interval includes the null value, a conventional null hypothesis test would have led to nonrejection. If the confidence interval does not include the null value, the null hypothesis would have been rejected.

A final advantage of confidence intervals is that as sample size increases, the confidence interval gets narrower, yielding a more precise estimate of the population parameter. On the other hand, as sample size increases, the area of nonrejection of the null hypothesis gets smaller. This makes it easier to reject the null hypothesis, but if researchers do reject the null, the area of uncertainty is larger, producing the paradoxical situation that a larger study yields less information.

Bayesian Inference

A third way in which probability can be used to help researchers decide about the truth of hypotheses is known as Bayesian statistics. As Cohen pointed out, the Bayesian approach asks about the probability that a particular hypothesis is correct, given the data observed, whereas traditional hypothesis testing asks what is the probability of the data, given the null hypothesis. Cohen argues that the former question is the question of interest.

Bayesian analysis requires that researchers specify a probability distribution under the research hypothesis, known as the *prior probability distribution*. They then collect some data and use the observed data distribution to calculate a revised probability distribution known as the *posterior probability distribution*. That is, they use the data to modify their belief about the true state of affairs. As Cohen also points out, this allows researchers to use successive studies to refine both their hypotheses and their parameter estimates. Unfortunately, Bayesian analysis is computationally more difficult and requires a better grasp of mathematics than do null hypothesis testing or confidence intervals, so these methods are not as widely used in the behavioral sciences.

Model Fitting

Another attractive alternative to null hypothesis testing is the fitting of specific models to the observed data. The question being asked is how well the model

fits the data, and alternative models can be compared by contrasting the degree of lack of fit. The entire area of structural equation modeling uses this approach.

The mean specified in a null hypothesis (μ_{NH}) postulates a specific model for each observation in a set of data. The *lack of fit* of the null mean for any given observation is given by the distance of the data point from μ_{NH}. The sum of squares for these lack-of-fit values (SS_{NH}) is the lack of fit for the null hypothesis model:

$$SS_{NH} = \sum (X - \mu_{NH})^2.$$

Likewise, the lack of fit for the sample mean is given by its sum-of-squared deviations:

$$SS_{Data} = \sum (X - M)^2.$$

From the principle of least squares, SS_{Data} must be less than SS_{NH} unless the two means are equal. The model-fitting approach asks whether the reduction in lack of fit is larger than researchers would expect to have occurred by chance. The approach generalizes to the case of two or more means, a regression line, or any other statistical model that one might want to fit. In each case, the question being asked is whether the more complex model (for example, using two group means rather than one combined mean) provides a reduction in lack of fit that is greater than a chance amount.

Robert M. Thorndike

See also Descriptive Statistics; Statistical Significance; *T* Scores

Further Readings

Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist*, *49*, 997–1003.

Kaplan, D. (2000). *Structural equation modeling: Foundations and extensions*. Thousand Oaks, CA: Sage.

Meehl, P. (1978). Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. *Journal of Consulting and Clinical Psychology*, *46*(4), 806–834.

Meehl, P. E. (1997). The problem is epistemology, not statistics: Replace significance tests by confidence intervals and quantify accuracy of risky numerical predictions. In L. Harlow, S. A. Mulaik, & J. H. Steiger (Eds.), *What if there were no significance tests?* (pp. 393–425). Mahwah, NJ: Lawrence Erlbaum.

Salkind, N. J. (2004). *Statistics for people who (think they) hate statistics* (2nd ed.). Thousand Oaks, CA: Sage.

Schmidt, F. L. (1996). Statistical significance testing and cumulative knowledge in psychology: Implications for the training of researchers. *Psychological Methods*, *1*, 115–129.

Sirkin, R. M. (2006). *Statistics for the social sciences* (3rd ed.). Thousand Oaks, CA: Sage.

Skorupski, W. P. (2007). Bayesian statistics. In N. J. Salkind (Ed.), *Encyclopedia of measurement and statistics*. Thousand Oaks, CA: Sage.

Thompson, B. (2006). *Foundations of behavioral statistics*. New York: Guilford.

INSTITUTIONAL REVIEW BOARDS

Institutional review boards (IRBs) are local universities given the responsibility of oversight for federally funded research involving human subjects conducted by members of the organization. Researchers from the local organization serve on the IRBs and often employ one or more staff members, including a compliance officer, to help facilitate the review process. IRBs' responsibilities include providing training so that researchers conduct research safely, approving research protocols that are designed to protect participants from harm, and making sure that potential human subjects are adequately informed of the risks and benefits of their participation so that they can give informed consent. IRBs must balance the risk of the research against its potential benefits in approving research protocols. IRBs have become an entrenched part of the research process in the United States. There is near-unanimous support for the overall goal of IRBs protecting human subjects from unnecessary harm.

Because the policies based on medical research are also applied to social science and humanities research, a number of controversies and problems resulting from unclear or changing definitions or from mission creep and excessive regulation of research have emerged as IRBs have increased in number, size, and scope. IRB critics believe that these concerns have a chilling effect on research; reduce and slow research productivity unnecessarily; and seem to give IRBs powerful control over researchers who feel they have little recourse. Supporters of IRBs see increasing regulation and enforcement as positive steps in further protecting human subjects from harm by overzealous researchers. Nevertheless, it is clear that scholars will continue to have to interact with IRBs as part of the research process.

A Brief History

IRBs grew out of concerns over serious human subject research abuses such as the experiments conducted by Nazi scientists during World War II. The Nuremberg Code issued by the War Crimes Tribunal in response to these atrocities is generally recognized as the first international code of medical research ethics. However, research abuses continued, such as in Tuskegee, Alabama, in the late 1940s when poor, Black men were denied available medical treatment for syphilis so that researchers could study the effects of the disease on them. Although most of the commonly cited research abuses involve medical research, behavioral or social science researchers also committed abuses. Perhaps the most commonly cited example is a set of psychological studies conducted by Stanley Milgram in the early 1960s. These studies would be considered unethical by today's research standards because the participants had been deceived into believing that they were giving excessive electric shocks to other people and were inadequately debriefed about the experiment.

In 1974, the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research was created and the National Research Act was passed by Congress. This law required the establishment of local IRBs to review and approve all human subject research that is federally funded. In 1979, the commission approved a guide for research with human subjects: *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research.*

Enforcement of the human research guidelines has fallen on a number of different federal agencies over the years. Currently, the Office of Human Research Protection (OHRP) in the U.S. Department of Health and Human Services oversees an estimated 3,000–5,000 IRBs across the nation that regulate medical, social, and behavioral science research.

Defining Human Subject Research

Because IRBs have responsibility over research involving human subjects, defining human subject research becomes important. In general, the researcher must directly interact with the human subjects for research to need IRB approval. Therefore, publicly available information, such as political speeches, published or broadcast information, or observations of public behavior in which individuals are not identified

(e.g., percentage of people using cell phones while walking at the mall), is generally not considered the purview of IRBs. However, IRBs may differ on how they classify visiting Internet chat rooms or other publicly available online forums.

The guidelines define research somewhat ambiguously as contributing to the body of knowledge. This has generally been interpreted to mean that class assignments do not need IRB approval because they contribute to individual learning and not to the general body of knowledge. Research intended for presentations at conferences or distribution in academic journals or books generally must be reviewed by IRBs. The ambiguous definition of research has created problems, particularly in journalism departments and departments involved in historical documentation such as oral histories. Most IRBs have concluded that journalistic activities do not need IRB approval, but practice may vary for oral histories.

Informed Consent

A critical part of the approval process for research protocols concerns informed consent. The IRB guidelines indicate that potential participants in a study must be informed of the purpose of the study, the procedures, potential risks and benefits of participation, and the consequences (if any) of withdrawing from the study prior to its completion. Potential participants must be allowed to ask questions about their participation before agreeing to participate and be made aware that their participation is voluntary and that they do not have to answer any questions if they do not want to and can withdraw from the study at any time.

Informed consent can be accomplished through an oral or written presentation of the information. Oral consent is recommended when a written consent form provides the only link between the human subject and the study.

Two common problems with informed consent involve complexity and length. It is recommended that consent be worded using eighth-grade reading levels in order to reasonably ensure comprehension. In an effort to be comprehensive, some consent forms become so long that human subjects fail to read them completely, in effect nullifying the benefit of being informed.

Finally, minors (under 18) cannot legally give informed consent under most circumstances; they can only assent to participate. A parent or guardian must give informed consent for the minor to participate,

and the minor then assents or agrees to participate. In the past, researchers could notify parents in a letter explaining that their children would be involved in a study and that they only needed to reply if they did not want their children to participate. Current policy does not allow for this form of "passive consent." Now, active consent must occur in which signed documentation from the parent must be received before a minor can assent to participate. This need for active consent has made it more difficult to conduct research on minors.

Levels of Review

In reviewing and approving research protocols, IRBs classify research into one of three levels of review. In general, studies in educational settings or using generally accepted research methods such as interviews and questionnaires are classified as *exempt* and can be reviewed and approved by an IRB staff member. Exempt protocols must still make sure that human subjects receive adequate information concerning risks and benefits in order to give informed consent, record responses in such a way that the names of participants are not identifiable, and expect that disclosure of the participants' responses would not put them at risk for civil or criminal liability. Studies classified as *expedited* must still involve only minimal risk to human subjects, but may involve more sensitive topics or unusual research methods. Expedited studies are reviewed and approved by one or more members of the board rather than staff members. Studies that involve more than minimal risk to human subjects or involve human subjects in a protected group, such as minors or incarcerated individuals, must receive a *full board* review. The majority of the board members present must approve the study.

Issues and Concerns

Although nearly all researchers agree with the general mission of IRBs of protecting human subjects, a number of issues or concerns have risen around IRBs. Each potentially hinders legitimate research.

Administrative Problems

A collection of narratives from active researchers published in a special issue of the *Journal of Applied Communication Research* in August 2005 edited by Michael W. Kramer and Debbie S. Dougherty provides a sample of the administrative problems perceived by researchers. Among the most prominent problems are slow responses or delays from IRBs and changing or incorrect guidelines being applied to protocols. Often, these are the result of applying policies and practices appropriate for high-risk medical research and drug testing to low-risk research in the social and behavioral sciences. This often results in multiple submissions to IRBs and further delays. These processing delays can frustrate faculty members, particularly those under the pressure of gaining promotion and tenure in a limited time frame. Compliance officers and board member cite changing rulings from OHRP, insufficient staff, and poorly written protocols as reasons for most delays.

"Mission Creep" or Excessive Regulatory Control

A number of concerns with IRBs were cogently summarized and addressed in *The Illinois White Paper*, *Improving the System for Protecting Human Subjects: Counteracting IRB "Mission Creep,"* published by The Center for Advanced Studies at the University of Illinois in Champaign. Mission creep occurs when IRBs begin to oversee activity that previously was not considered human subject research. Efforts by some IRBs to include journalistic activities would be a clear example of this. Mission creep also occurs through excessive regulation when research that should be classified as exempt becomes classified as expedited or full board, or when unnecessary precautions, those typically needed for high-risk medical research, are imposed on low-risk social science research.

Those disturbed by mission creep point to problems with IRBs interpreting the guidelines through changing or expanding definitions. For example, some IRBs consider typical research incentives (e.g., pay or extra credit for a class) as undue coercion. IRBs sometimes broaden the definition of risk to include experiences that are common to everyday life, such as recalling a mildly unpleasant event from the past. Excessive regulation occurs when all students are considered to belong to a protected group or when psychological counseling must be made available for research involving minimal risk. Supporters of the expanding purview of IRBs indicate that their thoroughness and care prevent potential problems.

Loss of Purpose

As IRBs have become more established, they have grown in size and budgets. A number of critics, such as medical researchers George Annas and Elliot Foucar, have suggested that IRBs no longer focus attention on protecting human subjects, and instead focus on protecting universities from lawsuits by disgruntled research subjects. Others suggest that the OHRP is more concerned with maintaining bureaucratic rules and regulations than protecting human subjects. As evidence of this, critics point out that in the vast majority of cases in which universities have had their research programs temporarily shut down for violations, there is no evidence that human subjects were harmed in any manner. In most instances, the shutdown is caused by missing or poor standard operating procedures, poor minute keeping, or lack of a quorum at the approval meeting. Supporters point to the lack of harm to human subjects as evidence of the success of the procedures.

Michael W. Kramer

See also Ethics and Research; Ethnography; Experimental Design; Field Experiments; Naturalistic Observation; Qualitative Research Methods; Quantitative Research Methods

Further Readings

Annas, G. J. (2001). Reforming informed consent to genetic research. *Journal of the American Medical Association*, *286*, 2326–2328.

Center for Advanced Studies. (2005, November). *The Illinois white paper: Improving the system for protecting human subjects: Counteracting IRB "mission creep."* Champaign: University of Illinois.

Dougherty, D. S., & Kramer, M. W. (2005). Organizational power and the institutional review board. *Journal of Applied Communication Research*, *33*, 277–284.

Dougherty, D. S., & Kramer, M. W. (2005). A rationale for scholarly examination of institutional review boards: A case study. *Journal of Applied Communication Research*, *33*, 183–188.

Foucar, E. (2002). How much oversight is necessary to protect human subjects? *Journal of the American Medical Association*, *287*, 716–717.

Hamilton, A. (2005). The development and operation of IRBs: Medical regulations and social science. *Journal of Applied Communication Research*, *33*, 189–293.

Milgram, S. (1974). *Obedience to authority: An experimental view*. New York: HarperPerennial.

Web Sites

Office of Human Research Protections: http://www.hhs.gov/ohrp

INSTRUCTIONAL OBJECTIVES

Instructional objectives are written by teachers or instructional designers and express the intended outcome of instruction. They specify what the learners will be able to do as a result of instruction by stating the performance standards that learners are expected to achieve. Instructional objectives also go by other labels, such as *behavioral objectives, performance objectives*, or simply *objectives.* Two examples of instructional objectives are as follows: (1) given two single-digit numbers, the students will be able to add the numbers without the aid of a calculator, and (2) given a diagram of an amoeba, the seventh-grade science students will be able to label the protoplasm, nucleus, cytoplasm, and the pseudopodia.

Instructional objectives do not state instructional procedures. Instructional procedures refer to what teachers do during instruction, including the media used and instructional activities. Instructional objectives, by contrast, are student-focused; these objectives center on what students should be able to do at the end of instruction. Thus, although instructional procedures are determined in light of objectives, the objectives state what the students should do rather than what the teacher should do. Instructional objectives are also different from instructional goals. A goal is a broad, general statement regarding the intended benefits of instruction. Typically, goals are long term, spanning entire units of instruction. Objectives, by contrast, are specific and proximal. Thus, the instructional objectives provide the level of specificity needed to guide daily activities and allow for monitoring of goal progress. A single instructional goal may be achieved after students have met a series of progressive instructional objectives.

How instructional objectives are conceptualized and written has been the focus of much work in educational psychology. The work of several researchers, such as Robert Gagne, Benjamin Bloom, and Robert Mager, has converged on a set of recommendations for writing objectives, and both behavioral and cognitive learning theories have established principles for the use of objectives in instruction.

Characteristics of Written Objectives

Instructional objectives are a critical piece of effective instruction because they clearly identify and communicate the intended outcomes for instructional events. When a teacher communicates the instructional objectives to students, he or she serves to clarify expectations and allows students to attend to the important aspects of instruction. Instructional objectives also provide the standards that students can use to evaluate their progress toward desired outcomes. For the teacher, instructional objectives inform the instructional activities that occur and the assessments used to evaluate student progress. Accordingly, instructional objectives are written prior to instruction during the planning stage, are used to determine the instructional procedures used during instruction, and provide the guidelines for practice- and assessment-related activities.

During the planning phase, teachers write the instructional objectives that will determine the course of instructional events. Good instructional objectives are characterized as specific, measurable, observable, and short term. Specific objectives are unambiguous and clearly state what the learner should be able to do. In writing objectives, teachers should avoid the use of words such as *know, understand*, and *analyze* because these words do not convey a specific, unambiguous action. Instead, objectives should be written with verbs that clearly state the observable action that students will be able to carry out at the end of instruction. For example, rather than stating that a student should "know" certain material, a good instructional objective may state that the student should be able to define, list, label, or match specific concepts from the material. Rather than stating that a student should be able to "analyze" the material, a good instructional objective may state that the student should be able to diagram, differentiate, question, or summarize the material.

In the study of instructional objectives, several lists of verbs have been developed that teachers can use when writing objectives. The most widely known of these is the list of verbs that accompanies *Bloom's Taxonomy of Instructional Objectives.* Bloom and colleagues wrote the taxonomy in the late 1950s as a means to organize objectives according to their cognitive complexity. The six levels of this taxonomy, ordered from the least complex to the most complex, are to know, comprehend, apply, analyze, synthesize, and evaluate. The lowest level of this taxonomy, to know, refers simply to the memorization and recitation of facts. Evaluation, however, not only requires factual knowledge but also the ability to make judgments about that knowledge. Each level can be translated into instructional objectives by using a set of verbs that corresponds to each level. The verbs describing knowledge and analysis that were listed earlier in this entry are two examples of these verb sets. Other examples include the verbs *compose, design*, and *expand* to identify clearly what is meant by synthesis. *Compare, contrast*, and *judge* are three verbs that can turn evaluate into a clear and unambiguous objective. Complete lists of these verb sets can be found in most references on instructional objectives.

In addition to the use of specific verbs, teachers must give consideration to several other factors when writing instructional objectives. For example, instructional objectives must be close in time and reference outcomes that can be both observed and measured. This means that objectives must describe a behavior that the teacher can either see or hear. A teacher can see, for example, students' written responses or answers to mathematics problems. The teacher can also see the results of students' science experiments or created representations. Musical performances, such as playing a scale or singing a particular note, provide good opportunities for objectives that are heard. Each of these shares the characteristic of specifying a behavior that the student must exhibit in some way to the external world.

The ABCDs of Written Objectives

Suitably written instructional objectives contain four elements that can be expressed as the *ABCDs of objectives*. In this system, A stands for *audience*, B for *behavior*, C for *conditions*, and D for *degree*. The audience is the intended group of learners to which the objective applies. The audience may be seventh-grade science students, first-grade readers, or fourth-grade learning disabled students. The benefit of attending to the audience when writing instructional objectives is that, by stating the learners to whom the objective applies, it structures the objective so that it is focused on the learner. Behavior refers to what the learners are to do and is indicated by the verb of the objective.

Conditions indicate the exact circumstances under which the behavior is to occur. Conditions clarify the resources that learners will have available at the time of the performance and the circumstances under

which the behavior must be performed. It is the condition element of the objective that specifies, for example, if the behavior must be exhibited independently or with the aid of a partner, if students will be permitted to use notes or books, or if mathematics problems can be solved with the aid of a calculator. The final element, the degree, identifies the standards for an acceptable performance. It is the degree element of the objective that clarifies the quality or quantity of the behavior necessary to achieve a level that is deemed to be sufficient. Degrees might be expressed as the percentage of problems to be answered correctly, the number of errors that can be made, or the amount of time that a behavior should require. The degree element is most difficult to express when the objectives correspond to subjective performances such as writing a good summary or creating a novel design. In these cases, it is important that the objectives are tied to clearly stated scoring criteria. A good summary, for example, may contain the main idea, three supporting details, and a synthesis statement. Sharing these criteria with learners may clarify expectations and further specify the stated objective.

Categories of Written Objectives

To write instructional objectives, teachers must also consider the type of behavior that is the target of the objective. Objectives can be categorized as psychomotor, cognitive, or affective. Psychomotor objectives refer to behaviors that the learner is to perform. Playing a musical instrument or executing an athletic move is an example of a psychomotor objective. In academic domains, psychomotor objectives can refer to the execution of activities such as properly arranging manipulatives, operating equipment, or managing behavior. Cognitive objectives refer to what the students will learn or the intellectual capacity they will acquire through instruction. There are various types of cognitive objectives, such as comprehension or problem solving. Most resources on instructional objectives recommend using the levels of Bloom's taxonomy to conceptualize cognitive objectives. Corresponding verbs can, of course, be used when writing these objectives.

Finally, instructional objectives can refer to affective outcomes. Affective objectives include students' attitudes, values, and expectancies. An affective objective may state, for example, that students in a mathematics class will have confidence in their ability to complete mathematics problems or that the students will understand the value of a particular mathematical procedure to their everyday lives. In general, affective objectives are the most difficult to both write and assess because it is difficult to determine the objective and observable behaviors that correspond to the desired outcome.

The Role of Objectives in Instruction and Assessment

Once written, objectives provide guidance for the design of instructional materials and activities. Although there is not a one-to-one correspondence between a written objective and the instructional methods that can be used to achieve the objective, a well-written objective does guide the teacher's decision-making process. When selecting instructional materials, for example, the teacher must choose materials that contain the necessary information and that present the information in a manner consistent with the objectives. In class, instructional activities should also provide opportunities for practice and feedback on the desired behaviors. Scaffolds to support student progress toward the intended outcomes may also be included in the materials.

Furthermore, because instructional objectives require an observable behavior, they are also used to select the assessments that are used to determine if the objective has been achieved. The in-class activities, take-home assignments, and class tests are all consistent with the stated objective so that the performances for which students are held accountable are the same as those communicated to them through the objective. Student performance on these assessments provides the opportunity for teachers and students alike to assess progress toward the intended outcomes. When final assessments, such as a unit exam, are given, the assessment should align with the objectives that preceded it. The conditions of the exam, for example, should be consistent with the conditions of earlier objectives. In short, students should be assessed only on behaviors and abilities that were part of earlier objectives.

Theories of Learning and Instruction

Several theories of learning and instruction discuss the role of instructional objectives in student learning.

Contingency contracts are recommended by behavioral learning theories as a method for the teacher and student to discuss and establish the goals and conditions for learning-related outcomes. Contingency contracts are used with individual students. In these contracts, the instructional objective serves as the statement of the desired terminal behavior. Conditions for the behavior and consequences for the outcome are also included.

Gagne recommends that objectives be incorporated into the instructional design process. In this theory, objectives are classified according to one of five categories of learning outcomes (e.g., verbal information, cognitive strategies). Each category corresponds to a set of critical learning conditions, and the conditions specify the environmental conditions the learner needs in order to achieve the goal. By considering the conditions of learning alongside the categories of objectives, the instructional designer can translate the instructional objective into the instructional design.

Peggy N. Van Meter

See also Bloom's Taxonomy of Educational Objectives; Contingency Contracts; Goals; Grading

Further Readings

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: Handbook I, cognitive domain.* New York: McKay.

Gagne, R. M., Briggs, L. J., & Wager, W. (1992). *Principles of instructional design* (4th ed.). Fort Worth, TX: Harcourt Brace.

Mager, R. F. (1984). *Preparing instructional objectives* (2nd ed.). Belmont, CA: Lake.

INTELLIGENCE AND INTELLECTUAL DEVELOPMENT

Modern Conceptions of Intelligence

The modern history of theory and research on intelligence has its twists and turns largely due to the unwieldy nature of the concept. On one hand, intelligence is a concept that has high currency, a valued human resource that people try to cultivate or harness for advancing their causes and agendas, individually or collectively. In the meantime, it is also an abstract, elusive concept with many faces. When a task force put together by the American Psychological Association (APA) reported "knowns and unknowns" about intelligence, it was not able to come up with a uniformly agreed-upon definition of what intelligence is, other than the following statements by Ulrich Neisser:

> Individuals differ from each other in their ability to understand complex ideas, to adapt effectively to the environment, to learn from experience, to engage in various forms of reasoning, to overcome obstacles by taking thought. Although these individual differences can be substantial, they are never entirely consistent: A given person's intellectual performance will vary on different occasions, in different domains, as judged by different criteria. Concepts of "intelligence" are attempts to clarify and organize this complex set of phenomena. (Neisser et al., 1996, p. 77)

Three observations can be made about the statements. First, there is some consensus in the research community on what are typically seen as constituents of intelligence or intelligent behaviors; however, there is no clear answer as to whether they are closely connected facets or just concepts loosely coupled together. Second, the statements treat intelligence squarely as a differential or individual difference concept, yet give much leeway for intraindividual and contextual variability. Finally, the statements highlight the term *intelligence* as a psychological construct, a conceptual tool conjured up by scientists to sort and organize observations at an abstract level, not a physical reality like height or weight.

Although Francis Galton started the tradition of research on psychometric intelligence, Alfred Binet and Charles Spearman were the two most prominent early pioneers of modern theory and measurement of intelligence. Binet, in collaboration with his doctoral student Theodore Simon, developed the first modern intelligence test for the purposes of identifying and helping children with severe learning difficulties in school. Although the purposes were to develop a more reliable and objective assessment than informal clinical observations could offer, Binet saw the instrument as a clinical tool for diagnostic and instructional purposes. He also believed that intelligent performance and behavior involves a set of processes that can be

identified in children's performance and targeted for intervention. In comparison with Binet, Spearman was a different kind of researcher. He firmly believed that human intelligence can be clearly defined and measured with accuracy. He used simple measures of sensory discrimination as an indicator of intelligence, believing that sensitivity to subtle differences and relations best characterizes intelligent persons. In contrast to Binet's inclinations as a clinician, Spearman was a mathematician, who preferred numbers to immediate observations. By parsing performance indices into a shared variance and residue variance, he formulated a two-factor theory: individual differences in intelligence can be represented as consisting of a general factor (g) and specific factor (s). When Spearman made a bold claim with a title of "'General Intelligence' Objectively Determined and Measured" for his famous 1904 article, Binet was not convinced; he doubted whether a phenomenon as complex as intelligence can be reduced to a single number or a set of numbers. Indeed, he made the counterargument that two individuals who obtain the same score might well use quite different skill sets. Binet was more intrigued by subtle individual differences observed during performance than the apparent simplicity of the mathematical solution offered by Spearman. This tension, revealed in exchanges between Spearman and Binet and alluded to in the quoted statements above by the APA task force, has lingered to date.

Intelligence: Structural or Functional?

Conceptualizing intelligence as a structural feature of mind starts with Galton, who, along with many of his contemporaries, viewed intelligence as a heritable mental faculty. The structural view of intelligence was reinforced by massive intelligence testing and consolidated by the then-newly invented correlation and factor analytical technique. Efforts to delineate psychometric intelligence culminated with J. Paul Guilford's mapping of various configurations of abilities based on content, process, and product, and John Carroll's reanalysis of hundreds of psychometric studies of human abilities. Although the structural view of intelligence implicitly assumes that intelligence is a capacity of some sort, some voices resist such reification. Some scholars argued that psychometrically measured intellectual performance is better seen as an index of the effectiveness and efficiency of mind

vis-à-vis an array of task conditions rather than mental entities. An alternative, functional account involves an understanding of the context in which performance is observed and assessed, as well as how the person carries out the task. Whereas the structural view defines intelligence squarely as a person characteristic, a functional view defines intelligence at the interface of an enactive person and an impinging environment, as fit execution of behavior or performance in that context. In short, when the structural view sees competence as a personal trait, the functional view sees competence as context dependent.

Intelligence: Nature or Nurture?

The early pioneers of intelligence research differed with respect to whether mental capacity is heritable. Whereas Galton firmly believed that intelligence is largely a heritable quality, Binet considered it a human condition that can be modified through education and social interventions. Whether the quality called "intelligence" can be improved through education is still controversial today. There is a pervasive pessimistic belief that people cannot do much when it comes to intelligence. Supporting evidence comes from twin studies that show that as one reaches adolescence and adulthood, genetics seems to play an even more significant role in one's intellectual performance than when one is younger. On the other hand, evidence also indicates that intelligence is a malleable quality, and schooling and effective instruction make a difference in one's intellectual performance and cognitive organization of personal experience.

Raymond Cattell developed a more differentiated scheme in which fluid intelligence (Gf), the ability to manipulate complex information and detect patterns and relations, is interpreted to have direct biological underpinnings, whereas crystallized intelligence (Gc) reflects the cumulative effect of experience and education. Thus, Gf represents true genetically based intelligence and Gc is simply a derivative outcome of Gf acting upon experience and knowledge. There is a body of research on cognitive aging that seems to support the notion that Gf tends to decline with aging, but this trend is compensated for by the incremental changes in Gc. However, research also demonstrates that environment (likely including education) produces greater effects on Gf than Gc, affirming an opposite contention that fluid abilities are among the most important products of education and experience.

Indeed, an important task not fulfilled by the traditional IQ tests is how to design instructionally relevant useful tests of fluid abilities, showing how well students can transfer their learning to novel problems. However, when one examines what carries the most weight in general intelligence and in predicting future performance, crystallized intelligence seems to be a winner. Thus, the contention that Gf represents "natural ability" and Gc is simply "achieved competence" is a controversial one.

The nature-nurture issue is also related to whether one views intelligence as ultimately a structural feature of the person, reflecting a capacity, or a relational and functional property, reflecting the person–task interaction. At least three major sources of intelligence can be identified: biological, domain experience, and reflective. Biologically based intelligence (e.g., neural efficiency) may be difficult to modify, but both domain experience and reflective thinking are subject to significant environmental influences.

Intelligence: General or Specific?

At face value, standard intelligence tests are a composite measure of performance on a variety of tasks, originally intended to be this way by Binet. The practical utility of such a test may lie in the very fact that it is a gross measurement, like the Dow Jones Index, sampling as many "active performers" as possible to obtain an overall estimate. The paradox is that the broader the range of tasks a test covers, the less psychologically meaningful the test becomes. The trade-off seems to be between specification of cognitive (and possibly motivational) processes given a cognitive task, and the breadth of representations of task conditions that affords performance consistency, stability over time, and predictive validity. Practicality aside, it becomes theoretically problematic when one attempts to interpret various factor structures as indicative of how the mind is structured. Spearman, for one, ventured into such a speculation when he interpreted the general factor as "mental energy." The three-stratum theory proposed by Carroll in 1993 represents an integration whereby cognitive abilities are represented as a continuum from the most general to the most specific. Yet many researchers questioned whether the factor structures can truly afford a theory of the structure of human abilities, or they just reflect a statistical artifact.

Nevertheless, there is still a strong belief among many students of intelligence that general intelligence exists and that standard intelligence (IQ) tests measure it well. When it comes to understanding the nature of general intelligence, some take a more reductionistic view by tracking it down to its neurobiological roots. Recent research of developmental biology also seems to suggest that prolonged thickening of cortices may underlie more advanced cognitive development of individuals whose IQ scores place them roughly at the top 3% of the population. Others characterize general intelligence as the ability to deal with cognitive complexity; that is, more complex tasks require more mental manipulations of information, hence higher demands on mental capacity. Intelligence tests have been found to be correlated with high-complexity tasks more than low-complexity tasks, although exceptions also exist. Research on mental retardation also suggests deficits in extracting abstract relations and principles.

Although psychometricians have tried to map out human abilities, general or specific, for a long time, Howard Gardner and Robert Sternberg brought to the field different breeds of theoretical perspectives and research evidence. Gardner views psychometrically measured IQ as representing a culturally narrow view of what is important for effective intellectual functioning. His theory of multiple intelligences, for good or ill, has successfully pluralized the concept of intelligence. Moreover, his argument that the mind is not an all-purpose information-processing device but is composed of many specialized modules dedicated to processing specific types of information was based on a set of neuropsychological evidence. The underlying argument that processes are always sensitive to content is now widely supported. In comparison, Sternberg's triarchic theory represents a more complex system of theoretical propositions, encompassing cognitive, experiential, and contextual dimensions.

The question of whether intelligence is unitary and general or pluralistic and domain-specific often carries a structuralist overtone concerning how the mind is innately structured and organized, regardless of environmental experiences. It does not take into account the possibility that mental capacities can also be shaped by task environments over time. Gardner expressed a structural view of intelligence when he conceptualized the mind as innately modular. His recent conceptions seem to move toward a more functional view. A functional view of intelligence is by nature more domain-specific and context-bound. It does not presume, however, that the mind is innately domain-specific.

Rather, domain-specific competence could be the product of adaptive efforts that might involve both domain-general and domain-relevant resources.

Early Conceptions of Intellectual Development

From the onset, developmental approaches to intelligence take a different approach from differential or individual difference approaches, assuming that it is a function of age-graded development; that is, sooner or later a person will go through similar developmental changes in his or her intellectual structures and functions. Although working in Simon's laboratory on child psychology, Jean Piaget saw something that was apparently missed by psychologists of his time; that is, children's thinking has their own "logic" that is different from adults'. Piaget envisioned a developmental process that is completely outside the radar of differential psychology: At different developmental levels, intelligence has its different structures and functions, showing different organizational principles in cognitive functions. Such a developmental orientation has turned out to be extremely fruitful in understanding how children act on and represent their world, and how educators can facilitate children's intellectual development. Oddly enough, although Piaget differed significantly from differential psychologists in his theoretical interests, his assumption about the existence of deep structures of intellect (schemes and operations) reflects a conviction similar to that of differential psychology; that is, mapping the structure of intellect is possible just as is mapping the physical structures of an organism.

However, Piaget also started off a tradition of developmental research that shows little interest in individual differences. The separate paths of developmental and differential psychology have yielded findings and theories that are virtually noncommunicative with each other. On one hand, developmental researchers are seeking an understanding of how an "average" child develops intellectually or cognitively over time, while regarding individual differences as trivial or "noises" to be dismissed. On the other hand, consistent differences in intellectual performance observed within any age have prompted differential psychologists to develop age norms for individual differences in academic and intellectual performance calibrated to months of age. Because IQ scores are age-normed, intellectual development from such a differential point of view is simply how stable these age-normed individual differences are

over time. Both traditions can be criticized as missing an important part of intellectual development. On the differential side, age-normed rank order scores mask incremental changes in intellectual functioning and qualitative changes or reorganization of cognitive functions. Indeed, age-normed standard scores help perpetuate the notion that intelligence is a fixed quality, and by a further leap of faith, a genetically based individual difference. On the normative developmental side, in seeking to understand a typical or "average" child at a given age, the normative developmental psychology is also guilty of neglecting vast individual differences in intellectual functioning and development, not only in terms of psychometrically measured individual differences, but also in terms of different pathways and trajectories.

Metatheoretical Assumptions Guiding Theory Building

There are three main metatheoretical frameworks guiding research and theorizing: mechanistic, organismic, and contextualist. Mechanistic perspectives would simply see intellectual development as a derivative outcome of changes in cognitive machinery (e.g., working memory capacity). Organismic perspectives would view intellectual development as regulated by internal rules of growth (e.g., disequilibrium or innate skeletal principles). Contextual perspectives would view intellectual development as fundamentally embedded in the broader sociocultural context and situated at the interface of person–environment interactions (e.g., zone of proximal development).

Whether development of intelligence should be considered differential, quantitatively or qualitatively, is a matter of whether the core ontological commitments endorse a differential provision. Mechanistic approaches can lead to theories of differential intellectual development, in reading or other cognitive development, as long as the development is traced to the basic cognitive structures and operations. Organismic principles can also yield insights into differential development of intellectual competence by specifying when individuals diverge in how they negotiate the maximizing of mastery of the environment on one hand, and the optimizing of one's affect on the other. The contextualist doctrine, by far, offers the most unconstrained versions of intellectual development. On one hand, it is liberating by permitting various environmental forces (including fortuitous events and life-changing encounters) and individuality (with all its idiosyncrasies) to play a role in intellectual

development. On the other hand, it opens up a multitude of developmental possibilities that may prove unwieldy for scientific research or may fall short of parsimony. Again, there is a tension on the nomothetic or universal versus ideographic or particularistic continuum where cognitive and intellectual development are concerned.

Notwithstanding the challenges of integrating differential and developmental theories and research in a theoretically coherent manner, developmental approaches stand the best chance to address the three basic questions regarding intelligence discussed earlier. Traditional intelligence testing, and its methodological backbone, factor analysis, represent a static method and therefore cannot afford insights into dynamic person–environment interactions and changes of intellectual functions over time. By studying the temporal sequences of person–task or person–environment transactions, the structural-functional tension can be addressed as an issue of how initial structure facilitates functionality and how functionality leads to new structures. By studying nature and nurture in reciprocation and interaction, how biological preparedness and environmental affordances and constraints work in concert as a system can be further explicated (e.g., passive, evocative, and active correlations between the person and the environment). By studying intellect in the making, it is possible to show contributions of both domain-specific and domain-general cognitive resources. Because individual differences in intellectual development are one of the most important educational considerations, it is imperative from an educational point of view to incorporate a differential provision of how individuals differ in their intellectual development, and how learning and instructional activities enhance intellectual functioning and development. New developments in the fields of developmental and differential psychology show good promise of such a theoretical integration and consilience.

New Developments on Intellectual Functioning and Development

There are five aspects of current thinking that facilitate an integrative approach to intellectual functioning and development:

1. The importance of functional and social contexts,

2. The reciprocal relationship between functioning (including learning) and development,

3. The role of nonintellective factors in intellectual functioning and development,

4. A life span perspective on intellectual development, and

5. Developmental processes and mechanisms leading to differential intellectual development.

The Role of Functional and Social Contexts

In a narrow sense of the term, *context* refers to a set of immediate conditions where specific intellectual performance is observed. It can include the nature of tasks involved, situations in which the performers find themselves, and instruments and criteria used to make observations and assessment as well as temporal changes of strategies and performance. This emphasis shifts from an exclusive focus on intellectual performance as a person variable (i.e., competence) to a focus on intellectual performance as relational property of person–task interaction in real time (i.e., performance). Microgenetic methods, which follow individuals' interactions with a specific task environment intensively for days and weeks, uncover intraindividual as well as interindividual variability in children's intellectual performance. Such an idiographic approach to studying real-time, micro-level developmental processes has significantly changed the way development is understood, from a view of monotonic incremental or structural changes to that of variation, selection, and optimization, not unlike the process of biological evolution. Such a micro-level process view of development integrates functioning and development.

At a more abstract level, context can mean different functional contexts where intelligent behavior may entail different sets of capabilities and propensities, such as academic versus practical settings. Thus, Brazilian children can perform well on "street math" but cannot do as well on equivalent "school math." Professional racetrack gamblers are capable of sophisticated reasoning in their domain of expertise but do not perform superbly on standard intelligence tests. Each domain or field may have different sets of abilities and propensities and different threshold requirements. Even within academics, requirements for successful adaptations may be quite different. An emphasis on context also implies a specific social milieu or cultural context where intellectual functioning and development take place. Lev Vygotsky insisted that all higher mental functions initially occur at the social level and then

are gradually internalized with practice. The role of more competent others is indispensable. The notion of zone of proximal development treats intellectual development as fundamentally mediated by socializing agents. The new movement on "situated cognition" also stresses the distributed nature of intellectual functioning.

The Relationship Between Functioning and Development

Current thinking and research involve a new understanding of relationships between functioning and development, and learning and development. Development is traditionally considered structural organization and reorganization of experience and cognitive and behavioral functions that evolve over time. Experience and knowledge facilitate this process but do not change its nature. Learning, on the other hand, is seen as a process of the acquisition of new knowledge and skills that has little bearing on the cognitive infrastructure. This view has been changed significantly. Research shows, for example, that children who are chess players performed meaningful chess memory tasks at a much higher level than adults who had not learned chess. The study demonstrates that knowledge significantly alters basic cognitive functioning. The emergent expertise literature provides compelling evidence that knowledge enables sophisticated reasoning and problem solving. Now, learning is seen an integral part of development. Integrating learning and knowledge into developmental theory entails a more functional, rather than structural, view of intellectual development; in other words, intellectual development is likely more contextual than organismic. It opens a new way of looking at intellectual development that stresses the importance of contextual experiences (including educational experience) and the facilitative role of others while incorporating internal principles governing the behavior of the organism at a specific developmental level.

Urie Bronfenbrenner and Stephen Ceci argued that biological potential for learning and intellectual development can be realized only through proximal processes, that is, transactional experiences with specific environmental contexts both immediate and mediated. Accordingly, learning potential in a given situation, which is typically considered an important aspect of intelligence, can be built into a differential theory of intellectual development. This is in keeping with Vygotsky's notion of zone of proximal development and its practical application, dynamic testing, which is not assessing what the person already knows, but how well the person learns in real time when given instructional hints and prompts. Dynamic testing helps diagnose a child's ability to learn at a specific level of competence, and it informs educational interventions aimed at assisting children with learning difficulties, reminiscent of what Binet was doing during the inception of intelligence theory and measurement.

The Role of Nonintellective Factors

Both psychometric theories of intelligence and age-graded normative theories of intellectual development tend to be ability-centric theories; that is, the kind of abilities a person can display in a perform-on-demand condition. Such ability-centric approaches elicit maximal performance at the expense of neglecting a person's typical engagement, which has much to do with personal dispositions, such as openness to experience. Kurt Fischer made a related distinction in developmental theory between functional-level and optimal-level development, an important revision of Piaget's theory, which assumes by default an optimal-level development (e.g., emphasizing what adolescents potentially can do, rather than what they are actually able to do).

Nonintellective factors may play what Bronfenbrenner called a development-instigative role. Intellectual dispositions, such as risk-taking and open-mindedness, also play an important role in how one approaches an intellectual challenge. Intellectual dispositions may lie in the intersection of personality and intelligence. It may be argued that nonintellective factors such as interest and persistence may be particularly important in differential intellectual development, as much knowledge and skill building entails sustained engagement and deliberate practice.

Personal agency in intellectual development can also take a more conscious form. Development is now seen as a self-modifying process, part of intellectual development as self-engendered changes. All of these conceptions involve a distinct role of the consciousness and self. One example is what Annette Karmiloff-Smith called *representational redescription* made by children to articulate otherwise implicit knowledge. Rather than seeing intellectual development as fundamentally cognitive reorganization due to changes in cognitive infrastructure or architecture, Karmiloff-Smith

sees the role of consciousness as reflecting a developing human tendency to exercise cognitive and metacognitive control over aspects of their environment. This is in line with a renewed emphasis on reflective thinking, which was stressed decades ago by Dewey as quintessential for effective intellectual functioning. Education seems to play an important role to enculturate such habits of mind.

A Life Span Perspective

Current thinking goes beyond adolescence to embrace a life span perspective on intellectual functioning and development. Biologically, the Gc and Gf distinction provides a useful perspective on when one reaches cognitive maturity and when cognitive or neural efficiency starts to decline with the aging process. However, intellectual development is constrained, but not dictated, by biological differences and changes; in other words, intellectual development is not as canalized as other aspects of development, such as language development. This point becomes clearer when one views intellectual development (Gf or Gc) as often enabled by learning experiences and education. It can also be argued that much of intellectual development occurs beyond adolescence. There is a growing body of research on adults' differential intellectual development. The findings of behavioral genetics research that identical twins are more alike when they reach and go beyond adolescence, and that nonshared environmental factors play a more distinct role in later development, can be incorporated and further tested in such research. The topic of intellectual functioning and development also overlaps with expert performance and development of expertise.

Coherent Accounts of Processes and Mechanisms

Attempts at integrating differential and developmental approaches have been made at the level of childhood, adolescence, and adulthood. However, an important task is to provide processes and mechanisms that account for differential development in a coherent manner. Some scholars suggest that mental development in the first 6 years is similar for everyone and then diverges due to both environmental and genetic forces. Others provide motivational principles explaining differential pathways and trajectories. Robert Siegler's overlapping-wave theory of cognitive

development also potentially explains how differential development occurs. Studies that incorporate both comparative and microgenetic methods may ultimately be capable of addressing the two central issues of intellectual functioning and development: a process account of development that entails the role of adaptive functioning, and a developmental theory that is capable of addressing differential pathways and trajectories. In this way, differences between nomothetic and idiographic approaches to understanding intelligence and intellectual development can be eventually resolved or reconciled.

David Yun Dai

See also Bell Curve; Emotional Intelligence; Fluid Intelligence; Gifted and Talented Students; Metacognition and Learning; Multiple Intelligences; Triarchic Theory of Intelligence; Vygotsky's Cultural-Historical Theory of Development

Further Readings

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies.* Cambridge, UK: Cambridge University Press.

Ceci, S. J. (1996). *On intelligence: A bio-ecological treatise on intellectual development* (2nd ed.). Cambridge, MA: Harvard University Press.

Ericsson, K. A., Charness, N., Feltovich, P. J., & Hoffman, R. R. (Eds.). (2006). *The Cambridge handbook of expertise and expert performance.* New York: Cambridge University Press.

Gardner, H. (2003). Three distinct meanings of intelligence. In R. J. Sternberg, J. Lautrey, & T. I. Lubert (Eds.), *Models of intelligence: International perspectives* (pp. 43–54). Washington, DC: American Psychological Association.

Guilford, J. P. (1967). *The nature of human intelligence.* New York: McGraw-Hill.

Lohman, D. F., & Rocklin, T. (1995). Current and recurrent issues in the assessment of intelligence and personality. In D. H. Saklofske & M. Zeidner (Eds.), *International handbook of personality and intelligence* (pp. 447–474). New York: Plenum.

Messick, S. (1992). Multiple intelligences or multilevel intelligence? Selective emphasis on distinctive properties of hierarchy: On Gardner's Frames of Mind and Sternberg's Beyond IQ in the context of theory and research on the structure of human abilities. *Psychological Inquiry, 3*, 365–384.

Neisser, U., Boodoo, G., Bouchard, T. J., Boykin, A. W., Brody, N., Ceci, S. J., et al. (1996). Intelligence: Knowns and unknowns. *American Psychologist, 51*, 77–101.

Perkins, D., & Ritchhart, R. (2004). When is good thinking? In D. Y. Dai & R. J. Sternberg (Eds.), *Motivation, emotion, and cognition: Integrative perspectives on intellectual functioning and development* (pp. 351–384). Mahwah, NJ: Lawrence Erlbaum.

Sternberg, R. J. (Ed.). (2000). *Handbook of intelligence.* Cambridge, UK: Cambridge University Press.

INTELLIGENCE QUOTIENT (IQ)

Intelligence is a construct that has been proposed by psychologists to underlie much of human behavior and is a significant factor contributing to an individual's ability to do some things more or less well. Most would agree that some children are better at math or language arts than others, or that some hockey players or musicians are gifted in comparison to their peers. It might be argued that some individuals are born that way, whereas others have the benefit of good environments and learning opportunities that can build on their basic abilities. The intelligence test, and resulting intelligence quotient or IQ, is a means for assessing and measuring intelligence, with the results often used to classify or select persons or predict such outcomes as school achievement.

Both the construct of intelligence and its measurement are not new, and both existed well before the advent of psychological science. Historians have traced the forerunner of current cognitive ability and achievement assessment to more than 2000 years B.C. Although intelligence has been studied in a number of ways, from an early emphasis on sensory processes to the more current attention given to brain-imaging techniques, the mainstay in the study and assessment of intelligence has been the IQ test. Psychologists not only assess intelligence but also study how intelligence is expressed; what causes it; and how it contributes to understanding, explaining, predicting, and even changing human behavior.

Despite intelligence being a much studied area of psychology, there is still considerable controversy and emotion regarding the use of the IQ and intelligence tests and the results gleaned from them in such contexts as schools and industry to describe both individuals and groups. Given continued advances in the theories of intelligence and cognitive assessment instruments, the issue appears to be less with the constructs and the tests used to measure it, and more with how this information is or can be used.

Theories of Intelligence

Psychology joined the scientific community in the late 1800s, and since then, a number of theories outlining human intelligence, accompanied by a huge body of research, have emerged. The hallmark of science and scientific inquiry is the creation of theories and the pursuit of empirical support for the hypotheses that are generated by and from a particular theory. The current theories of intelligence attempt to explain what it is, what causes it, and what intelligence tells us about other human behaviors.

Although research has demonstrated that there is a considerable genetic component to intelligence, it is also recognized that intelligence is an acquired ability that reflects opportunity and experience such as comes from effective schooling and home environments. Studies showing the remarkable similarity in measured ability between twins, whether reared together or apart, provide much evidence for a genetic foundation to intelligence. However, intelligence appears to be polygenic rather than located on a specific gene. Studies have also shown that animals raised in very restricted in contrast to "rich" environments not only show considerable differences in, say, their capacity to solve problems, but also show an impact on their brain structures (e.g., number of neural connections). As well, research has shown how the effects of poverty and restricted educational opportunities can negatively influence human development, including intelligence.

Among the environmental factors that are known to directly influence brain functioning and thus intellectual development and expression are various medical conditions, neurotoxins, drugs such as alcohol (certainly during pregnancy, as observed in children diagnosed with fetal alcohol syndrome), and chemical pollutants such as lead and mercury. Almost anything that negatively affects the brain, such as head injury and oxygen deprivation, will have small or large observed effects on intelligence and its expression. Less obvious but just as important are such additional factors as motivation, self-concept, and anxiety, all of which can influence a person's score on an IQ or intelligence test and their everyday functioning at work or school.

Culture also affects the expression of intelligence. Although there is a universal ability related to the capacity to acquire, store, retrieve, and use information from everyday experiences as well as from direct teaching (e.g., school), how this is expressed, the content of a person's response to a question, and the language used in providing an answer to a question all reflect the interaction between the person's genetic capacities and the environmental opportunities for intelligence. In addition, arguments have been made that what constitutes intelligence may vary across cultures and that different ethnic groups may have differing, but equally intelligent, reasoning strategies. On the other hand, the successful adaptation of contemporary assessment instruments for use in a large number of countries suggests that central abilities and capacities comprising intelligence may be shared across cultures.

It should also be pointed out that intelligence is also a developmental construct. A 5-year-old child has a very different view of, say, cause–effect relationships or the understanding of number concepts than does a 15-year-old in Grade 10 or a 35-year-old with a university degree or a 50-year-old working in a factory. Brain maturation very much influences the qualitative description of intelligence. At the same time, it has been demonstrated that intelligence does change across the life span, with some kinds of intelligence referred to as crystallized intelligence (e.g., a person's knowledge of words and language, learned skills such as solving arithmetic problems) more likely to remain unaffected and possibly continue to improve with age than are abilities reflecting fluid intelligence and speed of processing information (reflecting neural efficiency), barring, of course, dementia and other diseases underlying cognitive decline.

Another debate found in theoretical discussions and observed in models of intelligence is centered on whether intelligence is a single characteristic or is composed of several or even multiple abilities. These views can be traced back to the turn of the previous century, when psychologists such as Spearman argued that intelligence was a set of specific but related factors that resulted in an overarching general factor (essentially similar to the current full-scale IQ [FSIQ] score found on many tests). In contrast, Thurstone proposed that intelligence was made up of a number of primary mental abilities that could not be captured in a single summary score or an FSIQ.

Today's tests and models continue to reflect these divergent viewpoints. For example, psychologists such as Guilford have proposed that intelligence may have 120 or more facets, while Wechsler has argued for the relevance of the FSIQ (but also the importance of looking at both verbal and nonverbal performance). Other current models proposed by Sternberg, describe intelligence along the lines of practical, analytical, and creative abilities, whereas Gardner suggests that there are likely eight to nine core kinds of intelligence reflecting, for example, interpersonal intelligence (required for effective social interaction and communication), kinesthetic intelligence (observed in athletes who excel in their sport), musical ability (found in performers and composers), and logical-mathematical intelligence (reflecting the capacity to reason logically in mathematics and science such as physics). Other views, drawing from the work of Piaget, focus more on how intelligence develops (in stages) and how it can be encouraged through direct instruction and supportive learning environments (e.g., instrumental enrichment).

Thus, there is quite some diversity in how intelligence is defined, determining the key factors that affect its development and expression, and how it is best measured. Although this may be perplexing to some, it does show how complex intelligence is and, even more so, how very complex human behavior is. At the same time, a great deal is known about intelligence and what it tells us about human behavior. For example, intelligence tests, yielding a measure of general mental ability, are one of the best predictors of student achievement and success in elementary schools. On the other hand, and as expected, intelligence tests have been found to be more limited in predicting achievement among intellectually homogeneous populations. For example, university students generally possess average or above-average levels of intelligence such that divergent performance in this group appears to be more highly related to specific cognitive competencies (e.g., high aptitude in math) and personal attributes (e.g., motivation, study skills). Intelligence is additionally considered a key factor in understanding human capacity to manage stress and develop resiliency, psychological well-being, and even longevity.

History of Intelligence Testing

The very earliest tests of intelligence were not based on any particular scientific views and in many instances simply showed the wide or narrow range of performance on such tasks as strength of grip or pitch discrimination. More to the point, these tests did not tell us

about other human characteristics that, by expectation, they should. If intelligence is an underlying capacity that influences how well a person does in school, or a person's accuracy at solving arithmetic problems, or the speed at which he or she can perform other mental tasks, then the tests should be correlated with those behaviors and be able to predict how well a person may perform on those tasks requiring intelligence.

In contrast to the earliest tests of intelligence, more recent intelligence tests have resulted from extensive research efforts, while still garnering a great deal of misunderstanding from the general public. The first successful intelligence tests were developed by Binet and Simon at the turn of the last century and used in the schools of Paris, France, to help identify and classify schoolchildren according to their ability to learn and whether they would benefit from regular or special school programs. A short time later, these tests were introduced in the United States. Along with the Army Alpha and Beta intelligence tests used to screen military recruits during World War I, there was growing opinion that intelligence tests had considerable value for purposes ranging from personnel selection to identifying children who were intellectually gifted or retarded.

These early landmarks in the history of testing laid the foundation for the advancement and proliferation of subsequent intelligence tests. For example, the first intelligence test created by Wechsler, in 1939, has evolved into several recently published tests for assessing intelligence from preschool years to age 89, and these tests have now been adapted for use in a large number of countries. The number of tests available to psychologists for assessing cognitive abilities has grown considerably over the past 60 to 70 years.

Current Intelligence Tests

Today's intelligence tests vary from very brief measures that assess only a limited or narrow part of the broader intelligence framework (e.g., Raven's Matrices) to large comprehensive batteries that tap many different aspects of intelligence ranging from verbal comprehension and spatial reasoning ability to memory and processing speed (e.g., Woodcock-Johnson Cognitive). The large number of tests available also includes tests specific to various age ranges, both group and individually administered tests, brief and comprehensive batteries, and modified tests for use with, for example, hearing-impaired clients, or clients who are nonverbal or who are less proficient in English.

The majority of intelligence tests require a well-trained psychologist administering subtests that require the client to complete a range of tasks. Two broad types of tasks are used on intelligence tests—verbal and nonverbal. Verbal tasks generally entail a verbally presented prompt or question and require an oral response such as defining words (What is a *hammer?*), responding to general information questions (What is the distance between the *earth* and the *moon*?), or identifying similarities between two words (How are *convention* and *meeting* alike?). Nonverbal tasks usually involve visual stimuli or materials and/or require a psychomotor response such as copying geometric patterns using blocks, identifying important parts that are missing from both common and uncommon objects, or identifying patterns within a visual array. Although instructions and prompts to nonverbal tasks are sometimes given orally, verbal requirements are minimized within some tests through the use of gestures, modeling, or pictorial directions.

Both verbal and nonverbal tasks can be employed to measure a wide range of cognitive abilities and capabilities. For example, short-term memory may be assessed through a task requiring a student to repeat a string of presented numbers (verbal task) or to touch blocks in a previously observed order (nonverbal task). Regardless of the types of questions used, the psychologist is careful to ensure that administration and nonintellective factors do not confound the information gleaned from these tests. For example, it is necessary to make accommodations for persons with, for example, visual, auditory, or motor problems, lest these interfere with their performance on tests that should be specifically tapping intelligence.

The raw scores obtained on intelligence tests are given meaning by comparing them to the performance of large and appropriate reference groups. These group performance indicators, called *norms*, are based on extensive standardization studies whereby the test is administered to large numbers of examinees to both ensure that the test is working well and to build a comparison group that is similar on key characteristics such as age and that reflects the composition of the larger community (ethnicity, sex, socioeconomic status, etc.). An individual's raw scores on the parts and the whole test are then converted to standard scores through the use of tables; these standard scores are often referred to as IQ scores, and in the case of, say, the Wechsler Intelligence Scale for Children–Fourth Edition (WISC-IV), four index scores assessing verbal

comprehension, perceptual reasoning, working memory, and processing speed, together with a full-scale IQ, are reported with the average score set at 100. Furthermore, based on the use of normal curve proportions, scores of 130 would represent intellectually gifted persons; this score is obtained by less than 2% of the population. Scores of 115 suggest high average ability that exceeds the scores obtained by about 84% of the population. In contrast, scores of 85 are seen as low average; a person with this score would have scored higher than 16% of the population, but some 84% of the population scored higher than him or her.

For an intelligence test to be truly useful, it must demonstrate sound psychometric properties that include reliability and validity. For a test to be reliable, it should have a minimum of measurement error, thereby measuring with consistency and precision. Thus, a FSIQ score will have some error associated with it and should never been taken as an exact measure but rather one that reflects a range wherein the person's true score is likely to be. Validity means that the test in fact measures what it is intended to measure. If the test is supposed to measure acquired knowledge or crystallized intelligence, then it should do just that and not do something else. Although it can be said that current intelligence tests are among the very best measures used by psychologists, certain caveats still apply. For example, no one test tells everything about a person's full intellectual ability, because other factors, such as depression, low motivation, test anxiety, or cultural factors, can influence intelligence test scores.

Current Uses of Intelligence Tests

The use of IQ and other intelligence tests is a complex process that requires a comprehensive understanding and training in such areas as test principles (reliability, validity, test construction, norm groups, types of scores); human development; and test administration and interpretation. As such, certain state and provincial restrictions exist that limit who is permitted to administer and interpret the results. In general, the use of intelligence tests is limited to psychologists or other such individuals who have a minimum of graduate-level training in psychology and assessment.

Although most commonly used by school or clinical psychologists within school and clinical settings, intelligence tests may also be used by psychologists within other specializations (e.g., counseling, industrial organization, research) and in such additional settings as community and state agencies, workplaces, universities, and private practices. In part, the purpose for administering an intelligence test may vary to some extent depending on the reason for referral and who is administering it and in which setting. A school psychologist may use the results of an intelligence test to help decide which students should be selected for a gifted program, whereas a neuropsychologist may use the results to assist with determining the location and extent of a brain injury. In general, intelligence tests provide information that can inform a wide range of diagnostic and decision-making processes. Among the most common uses of intelligence tests are to assist with diagnostic and eligibility decisions, intervention planning, progress monitoring, and research into cognitive functioning.

Diagnostic and Eligibility Decisions

Originating with Binet and Simon's development of an intelligence scale to identify children who would benefit from regular and special education, one of the primary uses of intelligence tests has been and continues to be in making diagnostic and eligibility decisions. In particular, various classification systems, laws, and policies use an individual's level of intellectual functioning (IQ or equivalent) for the purposes of identifying particular groups of individuals and for determining eligibility for services. An example of this is the diagnostic criteria for mental retardation (MR) outlined within the American Psychiatric Association's *Diagnostic and Statistical Manual of Mental Disorders* (Fourth Edition, Text Revision) *(DSM–IV–TR)*, which assigns the severity of MR based on obtained IQ scores (see Table 1). Such diagnoses or classifications often assist individuals in accessing appropriate assistance from medical, educational, or mental health specialists and support an individual's eligibility for particular programs or services. Certain government agencies, for example, provide adults with IQs below a certain level with additional financial and individualized supports to participate more fully within society.

Whereas a global assessment of intellectual functioning is most commonly used for classifications requiring either low or high levels of cognitive functioning (e.g., mental retardation, giftedness), IQ scores within the normal ranges can also provide useful information for diagnostic and eligibility decisions. For example, the assessment of at least average cognitive abilities can support a diagnosis of a learning

Table 1 *Diagnostic and Statistical Manual of Mental Disorder* (Fourth Edition, Text Revision) *(DSM–IV–TR):* Specified Intellectual Impairment (IQ) Required for Assessing Severity of Mental Retardation (MR)

Severity of Mental Retardation	IQ Level
Mild	50–55 to approximately 70
Moderate	35–40 to 50–55
Severe	20–25 to 35–40
Profound	Below 20 or 25

disability when found in conjunction with particular cognitive processes that are compromised such that the child has considerable difficulty in the acquisition and use of certain skills (e.g., reading). Alternatively, an assessed absence of cognitive deficits can assist a psychologist in determining if noncognitive factors are contributing to an individual's poor performance in the classroom. By ruling out cognitive factors, the contributions of such additional factors as motivation or emotional problems can be explored further.

Intervention Planning

As with diagnostic and eligibility decisions, intelligence tests have a long history of being used to inform intervention planning. At the broadest level, identification of causes or particular diagnoses provides valuable information for treatment and intervention. For example, an individual identified as having severe mental retardation will likely require a program that provides extensive training in elementary self-care and social skills. Evaluation of intellectual functioning can also provide information about the general strategies and approaches to intervention. Individuals with higher levels of intelligence do well with indirect or inquiry-based instruction, whereas studies demonstrate the benefits of functional, systematic, carefully sequenced, direct instruction for individuals assessed with lower levels of general intelligence.

The link between a child's global IQ score and intervention is often indirect, however. For example, by knowing that a first-grade child is struggling in mathematics due to an ability problem, interventions may focus on practical activities to increase that child's counting skills and understanding of number concepts. In contrast, a child with motivational problems may demonstrate improvements if incentives for work completion are provided. It is also possible to more acutely identify appropriate intervention practices and remedial activities when specific cognitive abilities and processes are examined. In particular, the increased range of narrow and broad abilities assessed by contemporary intelligence tests provides an avenue for parsing out specific abilities, experiences, or learnings that may be contributing to depressed achievement or impaired cognitive performance. Results from current tests may suggest the need to target such areas as an individual's visual and auditory processing skills, short- and long-term memory stores, or processing speed. The greater specificity that is available from contemporary assessments is akin to the doctor who provides specific feedback regarding a patient's blood pressure, cholesterol level, and heart rate in the assessment of physical health. Significant problems identified in any one of these areas are associated with a prescribed course of action. Similarly, by identifying a particular cognitive ability that is underdeveloped, specific interventions can be implemented or compensatory supports put in place to ameliorate the identified weakness.

Monitoring Progress

Although an individual's cognitive functioning is considered to be fairly stable over time, changes in IQ scores can result from a number of factors, such as development, intervention, environment, and injury. As such, intelligence tests can be used to monitor an individual's intellectual abilities over time. For example, cognitive scores can be compared to determine the effectiveness of special educational programs, treatment (medications), and training. For individuals sustaining cognitive injuries, availability of assessment of premorbid cognitive functioning can assist with determining the extent of such injury and also monitoring the rate at which cognitive abilities are recovered.

Research

Of course, what is known about intelligence, in terms of what it is and how it contributes to understanding, explaining, predicting, and even changing human behavior has been heavily informed by intelligence tests. By studying the results of intelligence testing

within particular groups of individuals or across time, psychologists' understanding of what constitutes intelligence and the processes underpinning cognition deepens. For example, the use of intelligence tests with elderly populations has provided insight into how the brain ages and processes spared from more rapid cognitive decline. Similarly, psychologists' understanding of the basic psychological functions that are affected in certain disabilities and disorders is directly informed through the administration of cross-battery assessment approaches, in which tasks from a number of IQ tests are administered. Intelligence tests also feature prominently in research into the effectiveness of certain treatments or interventions.

Summary

Among the limitations cited against the use of intelligence tests includes possible error in measurement, curriculum insensitivity, bias against students of diverse linguistic and cultural backgrounds, and limited instructional applicability. Although merit and support can be found for arguments on both sides of such debate, recent research advancements and current "best" professional practices within the field of assessment serve to bolster arguments in favor of using intelligence tests in describing individual differences. First, the more comprehensive structure of human cognitive competencies and abilities assessed by contemporary intelligence tests greatly increases the utility of such instruments to better inform a wide range of diagnostic and decision-making processes. Second, it was once believed that intelligence is what the test measures, but it is currently acknowledged that contemporary intelligence tests yield information that only partially explains or accounts for "intelligence." There is additionally the recognition that test results need to be viewed in consideration of a myriad of additional factors (social, emotional, behavioral) and other information and knowledge about an individual (family structure). Psychologists' understanding of what constitutes intelligence and how it is developed and expressed is additionally enhanced through complementary avenues of investigation that use alternative methods (e.g., brain imaging techniques) or investigate aspects of intelligence previously unattended to (e.g., emotional intelligence). When contemporary assessment instruments are paired with multidimensional approaches to assessment, the test results, including those from reliable and valid

intelligence tests, serve as invaluable components in many diagnostic and decision-making processes.

Michelle A. Drefs and Donald H. Saklofske

See also Creativity; Individual Differences; Intelligence and Intellectual Development; Intelligence Tests; Multiple Intelligences

Further Readings

Birney, D. P., & Sternberg, R. J. (2006). Intelligence and cognitive abilities as competencies in development. In E. Bialystok & I. M. F. Craik (Eds.), *Lifespan cognition: Mechanisms of change* (pp. 315–330). New York: Oxford University Press.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies.* New York: Cambridge University Press.

Deary, I. (2001). *Intelligence: A very short introduction.* Oxford, UK: Oxford University Press.

Flanagan, D. P., & Harrison, P. L. (2005). *Contemporary intellectual assessment: Theories, tests, and issues* (2nd ed.). New York: Guilford.

Gardner, H. (2006). *Multiple intelligences: New horizons in theory and practice.* New York: Basic Books.

Gregory, J. R. (2007). *Psychological testing: History, principles, and applications* (5th ed.). New York: Pearson.

Kamphaus, R. W. (2005). *Clinical assessment of child and adolescent intelligence* (2nd ed.). New York: Springer.

Matthews, G., Zeiner, M. T., & Roberts, R. D. (2002). *Emotional intelligence and myth.* Cambridge: MIT Press.

Prifitera, A., Saklofske, D. H., & Weiss, L. G. (2005). *WISC-IV clinical use and interpretation: Scientist-practitioner perspectives.* New York: Academic Press.

Sattler, J. M. (2001). *Assessment of children: Cognitive applications* (4th ed.). San Diego, CA: Author.

INTELLIGENCE TESTS

An intelligence test is a structured situation designed to elicit information about the cognitive abilities of an individual. The test may be administered individually or in a group. There are certain advantages to having a test administered to a person one-on-one with a trained examiner, typically a school psychologist. No reading need be required on the part of the examinee, so it is possible to test young children and people of limited literacy. And an empathic and perceptive examiner can maintain a continuing high level of motivation on the part of the examinee.

Furthermore, the examiner can observe aspects of the examinee's behavior in the testing session that are not reflected in the numerical score but that would assist in diagnosis. On the other hand, individual testing is expensive—prohibitively so if information is desired on each child in a school or on each applicant for admission to an educational institution or a training program. Primarily because of cost considerations, most ability testing has become group testing, using paper-and-pencil or computer-administered testing instruments that can be scored objectively.

Scores on intelligence tests are normally distributed in large populations and are usually reported on a scale in which 100 indicates average intelligence. Scores are scaled so that about the top 16% of the population will receive scores of 115 or above, the top 2.5% will receive scores of 130 or above, the bottom 16% will receive scores of 85 or below, and the bottom 2.5% will receive scores of 70 or below. The majority of people—68% of the population—will have scores clustered around 100, ranging from 85 to 115.

The typical intelligence test will have a variety of items designed to tap different aspects of the person's cognitive abilities. Some of the items may ask for specific pieces of information, such as how many years there are in a decade or how much change a person would receive if he or she bought an article of clothing costing $18.67 and gave the clerk a $20 bill. Other questions might ask about objects missing or out of place in a picture; still others would be tests for memory, such as repeating back a list of five digits that has been read aloud, or tests of reasoning, such as finding the right pattern piece to complete a design. In an individually administered test, the examiner asks each question, records the answer, and makes a judgment as to the answer's correctness or quality. Testing stops when the examinee has failed to answer a specified number of questions correctly. When the test is administered to a group, the questions are often in multiple-choice format, and responses are usually recorded by filling in bubbles on an answer sheet. Answers are compared to a key, so judgment as to correctness is avoided.

Studies repeatedly reveal a link between scores on tests of intelligence and educational achievement. Those with higher test scores typically receive higher scores on standardized tests of academic achievement, earn higher grades in school, and complete more years of education. Research also shows a moderate association between intelligence test scores and job status and occupational success.

History

Alfred Binet is generally given credit for creating the first modern intelligence test in 1905. He developed it in France, where he was working in the public schools. In the 1908 version of his test, Binet introduced the idea of "mental level" as a way to express the cognitive ability of a child. The mental level of an item was the age at which the average child could solve that particular problem. An item that could be solved by the average child age 7 or above, but not by a child of age 6, was given a mental level of 7 years. Items were grouped by mental level, and testing ended at the first level where a child could not answer any items correctly.

Henry Goddard popularized Binet's 1908 test in the United States. Several English-language versions of the Binet scale were quickly developed by Goddard and others. In 1916, Louis Terman published an American edition that came to be called the Stanford-Binet and soon replaced all competitors. This test popularized the term *intelligence quotient* or IQ because scores were expressed as the ratio of mental level or mental age, divided by actual or chronological age. A child who tested "at age" received an IQ of 1.00. This ratio came to be multiplied by 100 to remove the decimal point, resulting in a scale where the average IQ is 100, the reference point still in use.

During World War I, American psychologists under the leadership of Robert Yerkes produced two new group-administered tests for screening Army draftees: a verbal form, Form Alpha, for those who could read English, and a nonverbal form, Form Beta, for the illiterate and those who did not speak English. Form Alpha had eight subtests and Form Beta had seven. Subtest scores were combined to produce a total IQ. After the war, the use of intelligence tests in schools, for college admissions, and in industry spread rapidly. Soon, several million tests, mostly of the group variety, were being administered each year.

In the 1930s, David Wechsler applied intelligence testing in the psychiatric clinic of Bellevue Hospital. The Stanford-Binet had been developed for use with children, and Wechsler needed a test for adults. He adapted for individual administration the tests from the Army testing program, grouped the tests into a verbal scale and a performance scale, and collected

norms for adults. The resulting test battery was called the Wechsler-Bellevue Intelligence Scale. The verbal subtests were combined to produce a verbal IQ, or VIQ, and the nonverbal subtests yielded a performance IQ, or PIQ. The verbal and performance scores were combined to produce an overall description of cognitive ability called full-scale IQ or FSIQ. At the same time, Wechsler rejected the IQ as the ratio of mental age to chronological age and introduced the scale that all intelligence tests use today, a scale with a mean of 100 and a standard deviation of 15.

Theories of Intelligence

Binet had viewed intelligence as a unitary concept that he equated with judgment. In 1904, Charles Spearman had suggested that all cognitive performances depended on a person's level of general cognitive ability, which he labeled *g*, and a specific ability required by that task and no other, called *s*. Spearman's "two-factor theory" (of *g* and *s*) agreed with Binet's ideas and his test, but it was rejected and harshly criticized by many American psychologists, particularly Edward L. Thorndike and his students.

A major factor in the debate was the nature of the battery of tests that each camp analyzed. Spearman usually had a small sample of subjects, each of whom had taken one test for each aspect of mental ability, much like those included in the Army testing program. Thorndike usually had a large sample of subjects who had taken a much larger set of tests, several for each type of ability. When L. L. Thurstone developed the methods of factor analysis in the 1930s and analyzed a very large battery of tests, he found that there were about eight identifiable dimensions of cognitive ability that he called the *primary mental abilities*. However, he also found that these primary mental abilities were all positively correlated. Spearman interpreted this finding that all mental abilities were positively correlated as vindication of his theory.

The debate over the structure of cognitive abilities continued through World War II. Thurstone's factor analytic results led to the development of batteries of tests to assess his ability factors as well as others. The most extreme of the multifactor theories was proposed by J. P. Guilford in the 1950s. Guilford's Structure of Intellect model eventually postulated 120 relatively independent abilities organized along the dimensions of Contents, Products, and Operations. Today, the main remaining parts of Guilford's theory are the concepts of divergent and convergent thinking. Divergent thinking is the ability to see nontraditional solutions to problems and is offered as a major component of creativity, whereas convergent thinking, the ability to extract a single correct solution to a problem, is seen as key to performance on most intelligence tests.

Cattell-Horn-Carroll Theory and Related Tests

In the 1940s, Raymond B. Cattell identified two broad classes of intellectual functioning, the ability to solve new problems, which he labeled *fluid intelligence*, or *Gf*, and the fund of knowledge and information one had acquired from experience, which he called *crystallized intelligence*, or *Gc*. Starting in the 1960s, Cattell and his student, John Horn, expanded this theory to include eight or nine broad abilities. A massive factor reanalysis of more than 400 previous studies led John Carroll to offer a similar theory in 1993. Carroll's three-stratum theory postulated about 70 narrow, specific abilities that could be grouped into 9 broad abilities, which in turn gave rise to a single general ability factor at the highest level. Cattell, Carroll, and Horn agreed that they had arrived at essentially the same place, and their unified theory is now referred to as the Cattell-Horn-Carroll or CHC theory. CHC theory now forms the theoretical basis for almost all of the commercially available intelligence tests.

Reaction-Time Theory

Beginning about 1980, a new line of research into intelligence began in the work of Earl Hunt and Arthur Jensen. A century earlier, Sir Francis Galton had proposed that reaction time could be used to measure intelligence, but the primitive instruments available lacked sufficient accuracy to detect differences. With the advent of computers capable of accurately timing both stimulus presentation and response time, speed of neural processing became a potent area of research on intelligence. As research using these methods progressed, speed of information processing (called the *chronometric approach to intelligence*) and the extent and efficiency of working memory were found to correlate substantially with scores on traditional intelligence tests. The massive body of work by Jensen has been particularly influential. He and his colleagues have shown that the aspect of traditional intelligence

tests that correlates most highly with processing speed/working memory is the single general ability factor most similar to Spearman's *g*.

One of the characteristics of processing speed that makes it attractive as a measure of intelligence is its relative freedom from cultural influences. One of the enduring criticisms of intelligence tests is that they are culturally biased, thereby resulting in lower scores for individuals who are not from the culture producing the test (generally upper-middle-class Whites in the United States). Critics of intelligence testing claim that the tests measure exposure to the majority culture and therefore underpredict the success of members of minority groups. Various lines of evidence do not support these claims, but the claims continue nonetheless. Because information-processing tasks can be designed to separate reaction time and movement speed from decision speed, relatively pure measures of an individual's ability to process information that are largely independent of experience are possible. The fact that these measures correlate most highly with those traditional intelligence measures that are least influenced by culture, such as Raven's Progressive Matrices, suggests that information processing and fluid intelligence are similar.

Contemporary Intelligence Tests

A large number of intelligence tests have been developed over the years to measure various aspects of cognitive ability. Those that are offered for sale are reviewed in the *Mental Measurements Yearbooks*, published about every 2 years by the Buros Institute for Mental Measurements at the University of Nebraska. Some, such as the Stanford-Binet Fifth Edition and the Woodcock-Johnson Third Edition, are intended for use with the full range of ages from 2 years to 80+, although the specific tests used at each age may differ. Others, such as the Wechsler Scales, have different tests for young children (the Wechsler Pre-School and Primary Scale of Intelligence), school-age children (the Wechsler Intelligence Scale for Children), and adults (the Wechsler Adult Intelligence Scale). Others, such as the Kaufman Ability Battery for Children, are only for a specific and limited age range, but almost all of them make an effort to measure some or all of the group factors in the CHC model. They all also offer a single index of general cognitive ability.

Educational Decision Making and Intelligence Tests

In educational settings, intelligence testing is used primarily to gain information to support a variety of placement decisions. School personnel need information to help predict how much an individual will learn from a particular educational program or how successful he or she might be in each of several alternative programs. For example, a high school may have to decide whether a freshman should be in the advanced placement section in mathematics or in the regular section. In addition, schools regularly must determine whether certain students meet eligibility critieria for special education services and, if they do, which services will best serve the needs of those students. In each of these cases, information helps the person making the classification or placement decision to identify the group to which an individual most likely or properly belongs.

A great deal of controversy exists surrounding placement decisions in educational settings. However, because of the existence of individual differences in achievement among students and the tendency for these differences to become larger, not smaller, as students progress through school, it is likely that such decisions will always be necessary. Students differ markedly in how much they have learned from previous instruction, their motivation to learn, the instructional conditions under which they learn best, their rate of learning, and their ability to apply previous experience to new learning situations. Furthermore, an instructional program that facilitates learning and growth for one student may hinder it for others. Classroom teachers and other educators have long recognized the need to adapt teaching methods, learning materials, and curricula to meet individual differences among students and to place students in the kind of learning environment that will optimize their educational opportunities.

Some students have instructional needs that differ so markedly from those of other students in the class that the teacher is unable to make the necessary adjustments within the classroom. Typically, these students are of two kinds. At one extreme are gifted students whose level of achievement and speed of learning greatly exceed those of other students in the class, and at the other are students who are having great difficulty mastering basic educational skills such as reading, are unable to adapt to classroom demands,

or are so disruptive that they interfere with the learning of other students in the classroom. A teacher may decide to refer these kinds of students to a school psychologist or counselor for assessment and possible alternative placement.

Assessment of Children With Learning Difficulties

Assessment of the construct of intelligence is central to the diagnosis of a number of prominent learning difficulties. Whenever a student is identified as a candidate for special education services, federal law requires that documentation, in the form of test scores, is collected and used as part of eligibility and placement decisions regarding that student. The type of test evidence needed to support different classifications varies by state and by classification, but a standardized test of intelligence is commonly among those administered. This is because intelligence tests are the most powerful tools employed in such evaluations. Because such tests are used diagnostically, a wide range of information can be gathered, ranging from the recall of general information to more processes. Although such tests are used to determine eligibility for special education, they are also very useful in the identification of the child's cognitive strengths and weaknesses—all valuable data when it comes time to provide a structured plan for the child's future learning-focused activities.

The classification of mental retardation almost always involves administration of an individual intelligence test such as the Stanford-Binet or the age-appropriate Wechsler Scale. Because limited intelligence is considered a principal feature of mental retardation, comparison of a student's test score to national norms of children of the same age is a necessary step in identifying this disability. An intelligence test score in the 70 to 75 range is one indicator of mental retardation but is not, in and of itself, enough evidence for a diagnosis. The individual must also show substantial deficits in adaptive and social functioning to be classified mentally retarded. Below-average intellectual functioning *and* an inability to adapt to the demands of normal life constitute mental retardation.

A diagnosis of a learning disability requires even more extensive assessments. Because learning disabilities can affect a child's (or an adult's) use of spoken or written language, performance of simple or complex mathematical calculations, and direct attention, there has to be a clear distinction between mental retardation and a learning disability. Intelligence tests are generally administered to ensure that a particular student's intellectual functioning is within the range of normal, and that the student is not delayed or mentally retarded. Standardized achievement can also be administered to help provide additional information regarding the teacher's in-classroom observations about a particular area of skill in an effort to more clearly identify specific deficits. Such observations can help better identify a discrepancy between intellectual capacity, as measured by the intelligence test, and level of academic performance. The discrepancy is the key to the identification of a learning disability.

Intelligence tests can be differentiated from achievement tests chiefly in the use to which the scores are put. Achievement tests measure the fund of knowledge one has acquired in a specific academic or occupational domain, such as language arts or mathematics. Achievement tests are postinstruction measures, and interpretation of the scores should normally be restricted to the domain and instructional experiences of interest. Together with intelligence tests, achievement tests help complete the picture of an individual's specific learning weaknesses.

For children with speech and language disabilities—one of the largest groups of children receiving special education services today—intellectual assessment is less important than for the other high-frequency disability categories, learning disabilities and mental retardation. The main reason is that it is entirely possible for a child to have difficulty in expressive ability, either oral or written, without affecting other, higher-order processing skills.

However, general tests of intellectual ability can provide useful information, especially for children who are mentally retarded or have a specific learning disability. Such tools can identify the role of psychological processes and allow the conclusion that multiple information-processing pathways are involved.

For children with disabilities that are low in frequency (e.g., visual or aural), assessment of intellectual abilities presents a challenge. This is because most tests are not suited, and not designed, for children with such disabilities. Consequently, testing must be conducted using modified formats with consideration of the limitations that the disability presents. The most significant concern when this is the case is that such accommodations are not used during the standardization process

and are often generated impromptu by the examiner. Not all learning problems are the result of disabilities. Some children are simply slower at developing certain skills, and learning difficulties may be just the result of maturational delays. Use of standardized tests of intelligence and academic achievement provide useful evidence to help tease these differences apart so that children receive appropriate educational experiences.

Criticism of Intelligence Testing in Education

It is often asserted that standardized tests of all types are used to deprive certain groups of access to educational opportunities. When tests are used in a mechanical fashion as in selection and placement activities, and due to the recognized fact that some groups in our society have historically performed less well on tests, standardized tests can unfairly bar members of underrepresented groups from educational opportunities. The term for this idea is bias in testing. Whether a test score or its use is judged to be biased depends, to a large degree, on the definition of bias being used. Given this concern, the question is whether individuals are indeed unfairly barred, and if this is the case, what then constitutes fair and equitable use of tests for the selection, placement, and classification of individuals?

There have been a number of well-documented court cases that have offered highly publicized decisions one way or the other on the issue of test bias. One such case specifically targeted intelligence tests. *Larry P. et al. v. Wilson Riles et al.* (1979) is the case in which the validity (an essential psychometric component of any test) of intelligence tests was called into question for Black children, in particular. The presiding judge ruled that standardized, individually administered tests of intelligence were biased against children who were culturally different from those on whom the test was normed. In a similar court case in Illinois, *Parents in Action on Special Education v. Hannon* (*PASE v. Hannon*, 1980), a different decision, with different implications, was reached: Standardized measures of intelligence taken in the context of a broader assessment system were found not to be biased against children from culturally different backgrounds.

The use and interpretation of test scores within groups whose cultures and experiences differ from the general population for which a test was designed have received a great deal of attention during the past 25 years. Although there are many subgroups in American society, ethnic and linguistic minorities are the most clear-cut of these and the ones for whom the appropriateness of tests and questionnaires is most open to question. Some critics contend that testing students from linguistic minorities in other than their preferred language produces inaccurately low estimates of true intellectual capacity. In *Guadalupe Organization v. Tempe Elementary School District* (1972), the court established that when the child's primary language is not English, the primary language the child does speak must be determined prior to assessment and that formal provisions to ensure accurate assessment must be made. Examples of provisions that can be made and were mentioned in this case include the administration by examiners who are fluent in English as well as the child's primary language, the presence of an interpreter to assist when necessary, the use of instruments that do not stress spoken language, and an assurance that the test results will be explained to parents or guardians in their primary language. The number of Americans for whom English is not the preferred language has increased such that many ability and achievement tests are available in a Spanish translation. Major test publishers now go to considerable lengths to ensure that their test items do not present an unfair challenge for students from different backgrounds.

Tracy Thorndike Christ and Robert M. Thorndike

See also Intelligence and Intellectual Development; Intelligence Quotient (IQ); Mental Age; Stanford–Binet Test; Working Memory

Further Readings

Flanagan, D. P., Genshaft, J. L., & Harrison, P. L. (1997). *Contemporary intellectual assessment: Theories, tests, and issues.* New York: Guilford.

Jensen, A. R. (1998). *The g factor: The science of mental ability.* Westport, CT: Praeger.

Thorndike, R. M. (1990). *A century of ability testing.* Chicago: Riverside.

Thorndike, R. M. (2005). *Measurement and evaluation in psychology and education* (7th ed.). Upper Saddle River, NJ: Prentice Hall.

Web Sites

Buros Institute for Mental Measurements: http://www.unl.edu/buros/bimm

Harcourt Assessment (also PsychCorp): http://www.harcourtassessment.com

National Association of School Psychologists:
    http://www.nasponline.org
Pearson's Assessment Group:
    http://www.pearsonassessments.com
Riverside Publishing: http://www.riverpub.com

INTERNAL VALIDITY

In experimental psychology research, it is important to confirm that results of a study are actually due to independent or manipulated variables, rather than extraneous variables. *Internal validity* is the degree to which the results are associated with the independent variable, and not other, uncontrolled factors. When an experiment is said to be internally valid, a direct causal relationship between the independent and dependent variables is demonstrated unequivocally. This is in contrast to *external validity*, which is the generalizabilty of results across different experimental settings. This entry outlines several types of extraneous variables that may jeopardize the internal validity of research in educational psychology.

Threats to Internal Validity

If an experiment extends over a long period of time, uncontrolled changes may occur in the experimental groups—this confound is known as the *history* effect. If a difference is observed between the experimental and control groups, it may be due to history effects or the interaction of history with treatment. For example, suppose an educational psychologist were to investigate the effectiveness of a reading program across the school year delivered to two groups (experimental and control) by specially trained teachers. Suppose further that for 2 months, a substitute teacher replaced the original teacher in one group—this may pose an uncontrolled effect, as the substitute teacher may have influenced students' reading performance.

As with history, *maturation* is also a concern when a study extends over a long period of time. As experiments go on, biological or psychological changes may occur in participants. For example, participants become older, and they may change emotionally or physically. For example, suppose an educational psychologist were to examine the effectiveness of a reading program over one school year. If factors such as physical, social, and intellectual development are not controlled for, these processes, rather than the reading program, may contribute partly or entirely to reading proficiency changes. Therefore, the investigator could not claim unequivocally that the reading program is beneficial; this study would lack internal validity.

Another threat to internal validity is *testing*. Many educational experiments use pretest and posttest designs; and often, both testing situations are similar. If so, then the participants may display a practice effect from repeated testing. If participants show an improvement on a particular test, it may be due to the repeat test, rather than the independent variable (such as the reading program). Additionally, subjects might be sensitized to certain aspects of the pretest on which they focus more, and, as a result, do better on the posttest. On the other hand, if different tests are used in pretest-posttest situations, this may cause differences in measurement that may bias the results. For example, if a pretest consists of a more difficult reading test and the posttest is easier, the improvement may not be due to the independent variable, the reading program, but to the differences in measurement. This is called an *instrumentation* effect.

Attrition occurs when there is loss of participants during the course of an experimental treatment. Participant dropout can bias the results because there may be important differences between the participants who discontinue in the study and the participants who remain in the study. That is, if certain types of participants are more likely to drop out of the study than others, then the group that concludes in the study consists of different types of individuals than the group that started. This can be detrimental to the results of experiments. If a study is measuring improvement to reading after the implementation of a reading program, and the students who drop out of the study are low-achieving students, then the results would be biased; the concluding sample would consist of high-achieving students and low-achieving students would not be represented.

Another threat to the internal validity of an experiment is *regression* toward the mean. When individuals score unusually low or high on pretests, posttest scores are likely to be less extreme and closer to the mean; over repeated testing, scores have a tendency to regress toward the mean (i.e., unusual scores are unlikely to occur for the same individual on both testing situations). If the scores on the posttest are different from those on the pretest, it may be due either to the independent variable or to regression effects.

When experimental studies use control and experimental groups, ideally the groups are equal on all things except the independent variable. However, if the groups are not equal, then *subject selection effects* may occur. As an example, suppose that an educational psychologist is interested in studying the effects of instructional technology in an introductory educational psychology course. The technology is implemented in the early morning class and not in the evening class. If the group that had used instructional technology performs better on an exam, the technology may not be the only factor that influenced this outcome. The two groups might be different in some aspects. Why would some students choose a morning section, whereas others choose an evening section? These potential differences in groups may contribute to the outcome of an experiment. Subject selection effects may interact with other threats to internal validity. For example, one group might mature differently and thus complicate the results; researchers would not be sure if the results of an experiment are due to the independent variable or to maturation effects. This threat to internal validity is referred to as *selection-maturation interaction.* Selection effects may interact with other threats to internal validity (e.g., a group might be affected by history or testing in different ways).

The final threat to internal validity is that of *experimental* or *expectancy bias.* When experimenters administrate tests and are in direct contact with the participants, they may inadvertently influence the results. If there is a change in the dependent variable in one group, it may be due either to the independent variable or to the influence the experimenter may have had. For example, individuals conducting the experiment may display unintended, differential treatment between the experimental and control groups. This may confound results.

Enhancing Internal Validity

Researchers who are aware of the potential problems that may arise in experimental research can create experiments that avoid or reduce the above-mentioned threats to internal validity. Randomized experiments are highly recommended as they can eliminate most of these threats (e.g., there is no subject selection bias if subjects are selected randomly, history effects will be reduced). Unfortunately, much of educational research compares treatments given to intact classrooms. Random assignment at the subject level is difficult, as

students are members of existing classrooms. In other words, classrooms cannot be split up easily, with only some children receiving various treatments.

It is also important to note that the presence of a threat to internal validity does not necessarily mean that the internal validity of an experiment is weakened. For example, maturation may occur in students between pre- and post-readings tests, yet not enough to invalidate the results of an experiment.

It is almost impossible to design the perfect experiment that completely safeguards all threats to internal validity. Rather, the internal validity of an experiment varies along a continuum. An awareness of these potential threats and ways of enhancing internal validity may protect an experiment from some of these problems.

Jennie K. Gill

See also Experimental Design; External Validity; Inferential Statistics; Quantitative Research Methods; Statistical Significance; Testing

Further Readings

Campbell, D. T. (1969). Reforms as experiments. *American Psychologist*, *24*, 409–429.
Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research on teaching. In N. L. Gage (Ed.), *Handbook of research on teaching* (pp. 171–246). Chicago: Rand McNally.
Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis for field settings.* Boston: Houghton Mifflin.

INTRINSIC VERSUS EXTRINSIC MOTIVATION

The concepts of intrinsic and extrinsic motivation are used in the fields of education and social psychology. In everyday language, intrinsic motivation is simply another way of saying that people are interested in and enjoy what they are doing. According to the social psychological literature, people are said to be intrinsically motivated when they do an activity for its own sake, not for any extrinsic reward. Extrinsically motivated behaviors are those in which an external controlling factor can be readily identified. If people solve puzzles, play games, or paint pictures for no obvious external reason, they are said to be intrinsically motivated. On

the other hand, students who study hard to obtain high grades, employees who work extra hours for pay, and children who do their homework to please their parents are said to be extrinsically motivated. This entry addresses the history of these concepts, theoretical perspectives, research related to intrinsic and extrinsic motivation, and the practical implications of the research findings.

Historical Overview

One of the earliest uses of the term *intrinsic motivation* came from primate research conducted by Harry Harlow and his colleagues in the 1950s. Intrinsic motivation was used to refer to the fact that monkeys would solve puzzles when solution to those puzzles did not result in extrinsic reward. Essentially, Harlow found that the animals performed efficiently "without resort to special or extrinsic incentives"; solving the puzzles seemed to be its own reward. In the 1950s, psychological accounts of motivation were dominated by Clark Hull's drive reduction and Sigmund Freud's psychoanalytic theories; both of these accounts emphasized primary drives (e.g., thirst, hunger, and sex). These theories could not easily explain Harlow's findings; the animals in Harlow's research were not motivated by primary drives. A review by Robert White in 1959 revealed additional evidence for non-drive-based motives such as exploration, curiosity, spontaneous activity, and manipulation of objects—motives that could account for findings such as those of Harlow. These new motives, subsumed under the term *effectance motivation*, form part of a process whereby an animal or person learns to deal effectively with the environment. Today, the term *intrinsic motivation* is used by many psychologists to refer to non-drive-based motivation.

One unresolved issue concerns the distinction between intrinsic and extrinsic motivation. The term *intrinsic motivation* is defined by the absence of obvious external factors such as extrinsic rewards. Although many human behaviors appear to occur in the absence of any obvious or apparent extrinsic consequence, they may, in fact, be a result of past consequences, or they may be due to anticipated future benefits. From this perspective, intrinsically motivated behavior is simply behavior for which appropriate controlling stimuli have yet to be specified. In other words, when we do not know why a person engages in a particular activity, we infer intrinsic motivation.

Thus, the motives for engaging in many activities get categorized as intrinsic motivation. As a result, behavior due to distant, hidden, or obscure external causes gets mistakenly labeled as intrinsically motivated. In spite of these conceptual difficulties, the terms *intrinsic motivation* and *extrinsic motivation* are frequently used in education and psychology.

Many theorists have been interested in the relationship between intrinsic and extrinsic motivation. Some theorists in organizational and industrial psychology assumed that intrinsic and extrinsic rewards were additive, combining to increase overall performance. In terms of work, the view was that the highest motivation to perform would occur when jobs were interesting and challenging and when employees received extrinsic rewards for performance (e.g., pay, recognition, promotion). That is, the belief was that productivity and satisfaction would be highest when intrinsic motivation was supplemented with extrinsic incentives.

In the 1960s, the psychologist Richard DeCharms challenged this view and suggested that intrinsic and extrinsic motivation may not add together, and that external rewards might actually interfere with or subtract from intrinsic motivation. DeCharms speculated that external rewards would change people's perceptions about the causes of their behavior. If people were rewarded for engaging in activities, they would begin to see themselves as doing the activity for the reward rather than for interest and enjoyment. In this way, DeCharms suggested, external rewards would undermine intrinsic motivation.

Research on Extrinsic Rewards and Intrinsic Motivation

The hypothesis that extrinsic rewards could disrupt an individual's intrinsic motivation has been highly influential in social psychology and education, leading many researchers to investigate the relationship between external rewards and intrinsic motivation. Two early experiments on the topic, one conducted by Edward Deci, the other by Mark Lepper and his associates, are frequently cited as examples of how extrinsic rewards negatively affect people's intrinsic motivation. Because incentives and reward systems are often used in business and education, the early findings on negative effects led to a splurge of research. Since the 1970s, dozens of experiments, using a common set of procedures, have been conducted to investigate the undermining effects of reward.

In a typical experiment, people are presented with an interesting task (e.g., solving puzzles, drawing pictures, or playing word games) for which they receive praise, money, candy, gold stars, and so forth. A control group performs the activity without receiving a reward. Both groups are then observed during a nonreward period in which they are free to continue performing the task or to engage in some alternative activity (free-choice period). The time that participants spend on the target activity during this period (free time), their performance on the task during the free-choice session, and/or their expressed interest in the activity are used as measures of intrinsic motivation. If participants in the reward condition spend less free time on the activity, perform at a lower level, or report less task interest than those in the nonreward group, reward is said to undermine intrinsic motivation.

Over the past few years, there have been several statistical reviews of experiments on rewards and intrinsic motivation. The most recent meta-analytic review by Canadian researcher Judy Cameron and her associates examined the results from 150 experiments. They found that extrinsic rewards do not always have negative effects on people's intrinsic interest; in fact, rewards were found to produce positive effects under several conditions. In general, on high-interest tasks, rewards were found to undermine intrinsic motivation when they were tangible, expected (promised beforehand), and loosely tied to performance. In contrast, tangible rewards increase intrinsic motivation when they are contingent upon achieving a specific level or standard of performance. Verbal praise and positive feedback also lead to positive effects on measures of intrinsic motivation. As well, all types of rewards have been found to increase intrinsic motivation on tasks of low initial interest. Recent research has shown that intrinsic motivation increases when rewards are offered and given for successfully achieving a graded level of performance leading to mastery. As well, when individuals are rewarded for achievement during learning and on tests, people show even higher levels of intrinsic motivation.

Theories of Extrinsic and Intrinsic Motivation

One theoretical account of the relationship between extrinsic and intrinsic motivation comes from Edward Deci and Richard Ryan's *cognitive evaluation theory* (CET). CET proposes that intrinsic motivation springs from two innate sources: the need for self-determination and the need for competence. Intrinsic motivation arises from activities that lead people to fulfill these basic needs. When people engage in activities such as puzzle solving, they are said to feel free and highly competent; these perceptions in turn activate the innate energy of intrinsic motivation. The energy of intrinsic motivation heightens interest and sustains involvement in an activity. Extrinsically motivated behaviors, on the other hand, refer to actions that are directly linked to an external cause such as an explicit reward, incentive, or threat. Extrinsically motivated actions are said to be characterized by pressure and tension and to result in a loss of perceived competence and personal freedom. These assumptions provide the theoretical basis for CET's account of the effects of extrinsic rewards on people's intrinsic motivation.

CET has typically focused on negative effects of rewards and their controlling versus informational aspects. From the CET perspective, all rewards are experienced as controlling, but rewards that are closely tied to achievement of performance standards are said to be the most controlling. Offering rewards for meeting performance standards leads individuals to feel pressured; such rewards undermine perceptions of self-determination leading to a reduction in intrinsic motivation. On the other hand, rewards linked to achievement can also provide information about competence. When people succeed at attaining a performance standard, the rewards convey information about performance that is positively evaluated by the individual; this evaluation leads to greater perceived competence and may offset some of the controlling aspects of rewards. For CET, the critical process for intrinsic motivation involves people's evaluation of rewards in terms of perceived self-determination, and the secondary process concerns perceived competence.

Another explanation of the effects of rewards on intrinsic motivation comes from attribution theory and Mark Lepper's *overjustification hypothesis.* As with CET, the focus has been on the negative effects of rewards; rewards are said to decrease intrinsic motivation by altering people's attributions of causation for their behavior. When rewards are offered for performing an activity, people are said to discount the internal causes of their actions (e.g., interest, effort, skill, ability) and to view their behavior as externally motivated. Heightened attributions of control by external factors lead to low levels of personal autonomy,

resulting in a loss of intrinsic motivation. At the same time, however, rewards for achievement may convey information about one's ability, mastery, or competence. Perceptions of high competence enhance intrinsic motivation. This is most likely when rewards are given for successful performance on an activity. Thus, rewards given for achievement may increase perceptions of competence as well as perceptions of external control; the effects of such rewards on intrinsic motivation will depend on the relative magnitude of these two competing processes. The overjustification hypothesis emphasizes the shift in causal attribution from internal to external sources as the basis for negative effects of rewards on intrinsic motivation. Perceived competence plays a role only if people attribute causation for performance to themselves (internal). Achievement cannot signify competence if people infer that their performance was caused by external factors (rewards or external pressure).

In contrast to social psychological theories that depict rewards as harmful to self-determination, competence, and internal attribution, Albert Bandura's *social cognitive theory* (also called *social learning theory*) holds that aspects of personal agency (e.g., self-motivation and self-directedness) may be enhanced by external influences such as rewards. That is, social cognitive theory rejects CET's appeal to innate sources of motivation and attribution theory's shift from internal to external sources as the bases for the effects of rewards on intrinsic motivation. Social cognitive theory contends that personal competencies (and other aspects of self-regulation) often are developed with the aid of extrinsic rewards.

Social cognitive theory deals with situations in which individuals extract personal standards from reward contingencies. The theory states that rewards given for achievement contribute to high personal standards. Attainment of high personal standards leads to task involvement (interest), positive evaluations of performance, and increased self-efficacy or competence. In contrast to CET and attribution theory, a high perception of competence or efficacy is the primary basis for intrinsic motivation—not self-determination or shifts in attribution. Another important distinction between social cognitive theory and other approaches is that achievement-based rewards can lead directly to greater involvement and interest. Social cognitive theory proposes that rewards for achieving challenging performance standards are likely to act as positive feedback leading to high task interest and involvement,

high personal evaluations of performance, and increased competence; these processes in turn increase intrinsic motivation.

Social cognitive theory distinguishes between non-competency-contingent and competency-contingent rewards. Non-competency-contingent rewards include rewards given without regard to mastery of performance (e.g., rewards offered for doing an activity with no criterion for performance). From a social cognitive perspective, non-competency-contingent rewards impart little indications of competency in that the rewards are loosely tied to behavior. This means that intrinsic motivation decreases when people are offered rewards simply for doing an activity without regard to level of performance, and meta-analytic findings support this view. Bandura describes rewards given for mastery (i.e., achieving relatively challenging behavioral standards) as competency-contingent rewards. Competency-contingent rewards are said to enhance intrinsic motivation by increasing people's perceptions of self-efficacy and competence. Cameron and colleagues found that intrinsic motivation increased when the rewards were tied to achieving mastery or a specific level of performance, a finding that supports social cognitive theory.

Practical Implications

Reviews of the research indicate that extrinsic rewards can have negative, neutral, or positive effects on people's intrinsic motivation. On tasks of low initial interest, extrinsic rewards can be used to increase motivation and performance. This finding indicates that rewards can be arranged to enhance time and performance on activities that initially hold little enjoyment. In education, a major goal is to instill motivation and enjoyment of academic activities. Many academic activities are not of high initial interest to students. An implication of this finding for education is that extrinsic rewards may be used by teachers to increase students' motivation and performance on low-interest academic activities.

On high-interest tasks, verbal praise and tangible rewards linked to success or to obtaining or exceeding a specific performance standard can enhance people's interest. These reward contingencies can be viewed as a subset of the many possible arrangements of the use of reward in everyday life. Extrinsic rewards can be arranged to progressively shape performance, to

cultivate initial interest in an activity, and to maintain or enhance effort and persistence at a task.

A negative effect occurs when an activity is of high initial interest, when the rewards are tangible and offered beforehand, and when the rewards are delivered without regard to success on the task or to any specified level of performance. How relevant to everyday life is the detrimental effect of reward, found in the experimental studies? In work settings, there are some circumstances in which individuals are rewarded irrespective of success or performance level. For example, because of compensation and promotion systems that are insensitive to performance (e.g., wage by job classification), some employees can vary their performance substantially with little effect on tangible reward. In such a situation, employees may have considerable latitude in how well or poorly they perform their jobs without any change in pay or fringe benefits. If an employee has high intrinsic interest in the job and receives a reward independent of performance, interest in the job may decrease, thereby producing poorer performance than if the reward were contingent on performance.

In educational settings, if a student has high interest in a particular subject matter, say, a history course, and receives the same grade regardless of performance, interest in history could deteriorate. The student may be less likely to spend time reading history during the course and following the conclusion of the course. Negative effects on intrinsic motivation also occur when rewards are promised but not delivered, offered in an authoritarian manner, or arranged for the benefit of the authority rather than the recipient. As well, rewards must be tailored to the individual; what is rewarding for one person may not be for another. When rewards are used inappropriately, people can actually perceive them as punishers rather than incentives, a point documented by the writer Alfie Kohn.

In order to avoid poor performance and reduced task interest in business and educational settings, employers, teachers, and administrators need to consider the basis upon which they allocate rewards, recognition, and advancement. Careful arrangement of rewards in educational settings and in the work environment can enhance students' and employees' interest and performance. This occurs when extrinsic rewards are closely tied to the attainment of performance standards. In addition, when tasks are challenging and graded in steps toward mastery of skills, extrinsic rewards given at each step of accomplishment can

instill interest and high personal standards for performance. As well, according to Robert Eisenberger's *learned industriousness theory*, rewards that are made contingent on effort lead people to work harder at an assigned task and show increased performance on other activities. Extrinsic rewards tied to high effort also result in higher intrinsic motivation than rewards given for low effort. Overall, rewards based on challenge, mastery, and high effort build skills, personal standards, persistence, and intrinsic motivation.

Current Directions

Research on extrinsic and intrinsic motivation continues to be important in psychology and education. One line of inquiry concerns the relationship between sources of motivation and creativity. Eisenberger and his colleagues have been working on the effects of giving rewards for novel performances. Creativity involves the generation of novel behavior that meets a standard or quality. Rewards for working hard at generating novel responses have been found to increase creative behavior. When rewards are given for creative thinking and performance on specific tasks, people show generalized creativity on other tasks. Rewards can also enhance creativity through increased intrinsic motivation. That is, rewards for high levels of creative performance have been found to increase perceived self-determination and competence, leading people to greater task enjoyment. In contrast, tying rewards to conventional-level performance (noncreative) causes people to be less creative and unmotivated.

Another line of research has focused on building intrinsic motivation without using extrinsic rewards. Deci and Ryan have developed self-determination theory as a more general approach to intrinsic motivation than CET. Self-determination theory focuses on the social conditions that activate or interfere with innate self-motivation. The theory postulates three psychological needs—competence, autonomy, and relatednes—that, when satisfied, lead to enhanced self-motivation. CET pointed to the importance of perceptions of autonomy and competence for intrinsic motivation; a third factor, relatedness, is also seen to enhance self-motivation. Relatedness concerns the social interaction dynamics between students and teachers, employers and employees, and parents and children. Secure relationships lead people to feel accepted, whereas negative interactions lead individuals to feel that others are cold and uncaring. Based

on relatedness, intrinsic motivation is enhanced or thwarted. Self-determination theory is being used in applied settings such as education, health care, parenting, work organizations, sports, and mental health.

Self-determination theory emphasizes the importance of social contexts for intrinsic motivation. Current research is showing how social contexts can modify the effects of extrinsic rewards on intrinsic motivation. Specifically, giving rewards for activities within social contexts that emphasize choice, self-monitoring of performance, and self-pacing have been found to increase people's intrinsic motivation. As well, rewards in these contexts lead people to perceive themselves as self-determined and competent. This research points to new approaches for arranging autonomy-supportive social contexts in classrooms and the workplace that promote enjoyment of activities using competency-contingent rewards.

Another direction of research concerns testing the major theories that attempt to account for how extrinsic rewards affect intrinsic motivation. As noted, CET, attribution theory, and social cognitive theory each emphasize different cognitive processes that mediate the effects of rewards on intrinsic motivation. CET emphasizes people's perceptions of self-determination and feelings of autonomy, attribution theory focuses on the shift in causal attribution from internal to external sources, and social cognitive theory postulates that perceptions of competence or self-efficacy are the major underlying processes. Recent research by Cameron and associates using path analysis has indicated that perceived competence is the central cognitive factor. When rewards for performance make people feel competent, intrinsic motivation is increased. The study also showed that rewards can activate processes that involve both internal and external sources of motivation. External and internal motivations add together to affect overall motivation for an activity, as suggested by early investigators in organizational and industrial psychology. Further research on cognitive processes that mediate reward effects on these two sources of motivation is warranted.

Judy Cameron and W. David Pierce

See also Cognitive View of Learning; Motivation; Self-Determination; Self-Efficacy; Social Learning Theory

Further Readings

Bandura, A. (1986). *Social foundations of thought & action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice Hall.

Cameron, J., & Pierce, W. D. (2002). *Rewards and intrinsic motivation: Resolving the controversy.* Westport, CT: Greenwood Press.

Eisenberger, R., & Cameron, J. (1996). The detrimental effects of reward: Myth or reality? *American Psychologist, 51,* 1153–1166.

Kohn, A. (1999). *Punished by rewards: The trouble with gold stars, incentive plans, A's, praise, and other bribes.* Boston: Houghton Mifflin.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55,* 68–78.

Sansone, C., & Harackiewicz, J. M. (Eds.). (2000). *Intrinsic and extrinsic motivation: The search for optimal motivation and performance.* San Diego, CA: Academic Press.

K

The most important human endeavor is the striving for morality in our actions. Our inner balance and even our very existence depend on it. Only morality in our actions can give beauty and dignity to life.

—Charles Dickens, *Dombey and Son*, 1847

KOHLBERG'S STAGES OF MORAL DEVELOPMENT

The word *moral* as it applies to human conduct and thinking has several commonly understood dictionary meanings. These include to act with virtue, to act honestly and with right character, and to act justly. The word *moral* is also defined as a generally accepted standard or custom for right living in a society. All of these meanings are embedded within Lawrence Kohlberg's stage theory of moral development.

Historically, the study of morality has fallen within the purview of theology and philosophy. Because of the influence of the positivist tradition in science, the study of moral development as a plausible subject of scientific inquiry did not fully emerge until the latter half of the 20th century. This breakthrough was significantly influenced by the challenges that Thomas Kuhn made concerning the assumptions of the positivist tradition. During this same period, cognitive theory gained increasing influence over behavioral theory toward its efforts in explaining reasoning development in children and adolescents. Cognitive developmental theory, through the pioneering works of Jean Piaget and Lawrence Kohlberg, has become pivotal in creating the study of moral reasoning development as a legitimate research area in educational psychology.

Cognitive Developmental Theory: Piaget's Work

Although Piaget is most noted for his theory of cognitive stage development, in 1932, he published a highly influential book titled *The Moral Judgment of the Child.* Piaget asserted that a child's moral reasoning develops not merely as a function of internalizing the norms and values of an individual culture, but rather as a natural process of constructing ideas of justice and fairness largely through peer interaction. Through his interviews with children, Piaget identified three phases of moral reasoning development. He considered these phases, which sequentially included the nonmoral, the heteronomous, and the autonomous, to describe only developmental trends. He noted that children move from an external morality in which justice is judged by concrete events toward a morality judged by pragmatic acts of reciprocity and finally toward a more internalized morality in which just acts are judged by context, intentions, and an idealized sense of reciprocity.

Kohlberg's Theory of Stages

Except for some notable differences in methodology, which included the addition of adolescents to his subject sample and the use of dilemma vignettes in place of story pairs in his interviews, Kohlberg's 1955

dissertation study was initially intended to be an empirical extension and validation of Piaget's work. Instead, his dissertation results became the foundation for Kohlberg's stage theory of moral development, which, in later studies, departed somewhat from Piaget.

First, Kohlberg argued that his subject responses should be categorized not merely in terms of overlapping phases but rather in terms of six distinct stages. These are further classified by three levels: (1) the preconventional, which includes the first two stages; (2) the conventional, consisting of stages 3 and 4; and (3) the postconventional, which includes stages 5 and 6.

Second, in contrast to Piaget's less assertive idea that moral development proceeds merely in a definite direction, Kohlberg claimed that his stages are hierarchical and progress in an invariant sequence, without stage reversal or stage skipping. Finally, whereas Piaget placed great emphasis on the importance of peer interaction, Kohlberg argued that social interactions between adults and children involving role taking of many kinds are important as well for the moral development of children and adolescents.

Dilemmas and Questions

Kohlberg's method of data collection consisted of presenting his subjects with hypothetical dilemma vignettes and then asking them a series of questions. The subjects' responses were his data. The following is an abbreviated version of one of his well-known dilemmas.

In Europe, a woman was near death from a cancer. There was one drug that doctors believed might save her, but the druggist was charging 10 times what the drug cost him to make. The sick woman's husband, Heinz, went to everyone he knew to borrow the required money, but he could only obtain half of what the drug cost. He told the druggist that his wife was dying and asked him to sell it cheaper or let him pay later, but the druggist said no. So Heinz got desperate and broke into the man's store to steal the drug for his wife.

In his interviews, Kohlberg would then ask his subjects questions such as the following: Should Heinz have stolen the drug? Which is worse, stealing or letting another person die a preventable death? What would you have done if you were Heinz?

Levels and Stages: Kohlberg's Criteria for Classifying Subject Responses

Kohlberg used three criteria for judging and classifying subjects' interview responses by level and stage. A response was judged for what a subject indicated was right or fair behavior, by the reasons given for right conduct, and by the indicated social perspective taken by the subject.

At the preconventional level, for stages 1 and 2, subjects' responses indicate a social perspective involving only the welfare of the self or someone identified with the self, such as Heinz. Also, right behavior is judged by its concrete pleasure and pain consequences to the self, and the reasons are relatively few in number and simple in content. Stage 1 responses indicate that right behavior involves the avoidance of punishment, so it would be all right for Heinz to steal the drug as long as he did not get caught and punished. Stage 2 responses are more advanced in that right behavior is viewed as not only serving one's own interests but also recognizing that others seek to serve their own interests as well. It would be all right for Heinz to steal the drug if he did not get caught and if he cared enough for his wife to take that risk. If Heinz did not care enough for his wife, then he should let her die.

At the conventional level, for both stages 3 and 4, the social perspective involves the acceptance of social and societal conventions for judging right and wrong. At stage 3, what is important is the welfare of relationships with important people, such as family members, friends, and schoolmates. Right behavior involves living up to the expectations of, and gaining approval from, those close to the self. Being good means maintaining mutuality in such relationships through trust, loyalty, respect, and care. Reasons for right conduct include that one should care for others and relationships in this sense and that rules that support this notion of good behavior should be followed. Despite the recognition that stealing and dishonesty are wrong, Heinz should still steal the drug because that is the level of caring that a husband should have for his wife.

At stage 4, the social perspective becomes more abstract in that the laws of society or the social order structure morality. Good citizenship to the state

becomes primary as a criterion for right conduct, which should be interpreted in terms of both the letter and the spirit of the laws of state. If Heinz had indeed exhausted all means of legally obtaining the needed money, then Heinz should steal the drug, but the courts should show clemency to him for the special circumstances surrounding his act of stealing.

Kohlberg argued that responses in his first three stages somewhat paralleled those of the three phases in Piaget's work but that at stage 4 and beyond, subjects used more abstract and sophisticated reasoning for justifying right conduct, which also reflects the highest Piagetian level of thinking and reasoning, that is, formal operational thought.

At the postconventional level, for both stages 5 and 6, the social perspective becomes more abstract and universal in that it is the welfare of all humankind that becomes critical for moral judgments. At stage 5, one is aware that universal human rights have priority over social conventions and laws but that the latter are necessary to structure morality within a social order. Thus, as in a social contract, one should always abide by conventional laws except in cases in which they conflict with universal human rights. However, stage 5 individuals sometimes have difficulty integrating the conflicts between legal and moral points of view. In such cases, the utilitarian principle of the greatest good for the greatest number should be followed.

Stage 6 responses, which are very rare, are based upon self-chosen ethical principles of justice. At this stage, individuals show the highest level of integration of thought, priority of values, and behavior with respect to moral issues. Right conduct is in accordance with universal principles, such as the respect for the dignity of all human beings as individuals and the quality of human rights. While recognizing the importance of the property rights of the druggist, Heinz should steal the drug, not merely because he is the husband of a suffering wife but because the preservation of life should have priority over property rights, and breaking the law would have the potential to improve the conventional legal system.

Validations, Controversies, and Criticisms

Several cross-cultural as well as longitudinal studies support Kohlberg's claim of developmental progression through the first four stages, but development beyond stage 4 is weakly supported at best. From his 1984 longitudinal results, Kohlberg reported that only 13% reached stage 5 and none of his subjects had clearly reached stage 6. This led to some revisions. Although he defended stage 5 as an empirically supported concept, Kohlberg suspended his empirical claims for stage 6 as an end point for moral maturity. Others have questioned if stage 5 is empirically valid for subjects in non-Western cultures because it is conceptualized in the philosophy of Immanuel Kant and other Western philosophers.

One of the most influential challenges to Kohlberg's assertion that his stages are hierarchical with respect to gender differences, stems from the work of Carol Gilligan. This is based on the fact that subjects in Kohlberg's early studies were exclusively male. In later studies, when he began to interview both adult male and female subjects, he found that females predominantly reasoned at stage 3, whereas males more frequently reasoned at stage 4. Noting that females are more likely to be socialized for adult leadership roles that require maintaining relationships through caring, Gilligan has argued that stage 3 should not be considered as less mature than stage 4 but rather only different and representing the impact of the typical differences in gender role socialization.

Educational Implications

Like John Dewey, Kohlberg believed that moral development and education were inextricably related. Based upon his research in the 1970s and 1980s, he developed an approach to moral education emphasizing the idea of building a just community in the classroom. He believed that moral dilemma discussion was the key process but that it required the following qualifications. Students and teachers engage in discussions that are challenging, and the dilemmas discussed are an integral part of the subject matter. Also, students and teachers discuss dilemmas from a mixture of different perspectives and stages of reasoning. Thus, despite controversy and criticism, Kohlberg, more than anyone, has had the greatest influence on the field of moral development study in relation to education, educational psychology, and developmental psychology over the past five decades.

Richard E. Hult

See also Moral Development; Social Development

Further Readings

Gilligan, C. (1982). *In a different voice: Psychological theory and women's development.* Cambridge, MA: Harvard University Press.

Kohlberg, L. (1981). *Essays on moral development: Vol. 1. The philosophy of moral development.* San Francisco: Harper & Row.

Kohlberg, L. (1984). *Essays on moral development: Vol. 2. The psychology of moral development.* San Francisco: Harper & Row.

Kohlberg, L., & Kramer, R. (1968). Continuities and discontinuities in childhood and adult moral development. *Human Development*, *12*, 93–120.

Kohlberg, L., & Mayer, R. (1972). Development as the aim of education. *Harvard Educational Review*, *42*, 449–496.

Piaget, J. (1965). *The moral judgment of the child* (M. Gabain, Trans.). New York: Free Press. (Original work published 1932)

L

I am learning all the time. The tombstone will be my diploma.

—Eartha Kitt

LANGUAGE DISORDERS

Language plays a central role in the conduct of human transactions. It is the vehicle by which we form interpersonal relationships. It is also a vehicle by which we gain access to knowledge and store the information that we've learned in memory. Finally, language is a means by which we create new knowledge, including great literary works. Indeed, some have argued that language is one of the defining characteristics of the human species. It comes about as an interaction of our genetic endowment and an environment that nurtures its emergence. Typically, language emerges effortlessly, but this is not always the case. When a child struggles to acquire his or her first language, with its onset delayed and development protracted in the absence of any other sensory or cognitive deficits and an intact environment, that child is considered to have a *language disorder* or *language impairment.*

Language is a multifaceted phenomena, with particular *form* through which unlimited *meaning* may be expressed and understood in the context of the situation of its *use.* A disruption may occur in any of these three facets of language or in their critical interactions, resulting in language disorders. Although the cause of such disruptions is typically unknown, the result is staggering because of the high societal value placed on verbal skill. Social interaction, knowledge acquisition, and one's very quality of life are all jeopardized by a failure in typical language development. Understanding language disorders is particularly important in the school setting, given its mandate for the academic development of children and the key role schooling plays in their socialization. Typically, the delayed onset of language is identified in the preschool years, yet the impact of a language disorder is felt well into the school years and beyond, making it important for early educators, classroom teachers, and other education professionals to be alert to and to understand the debilitating impact that a language disorder has on a child's life.

Basics of the Linguistic System

To understand language disorders, the multifaceted nature of the language system itself must be understood, because any or all aspects of it may be impaired. The language system is often thought of as arising out of the intersecting components of form, content, and use. The form of language comprises small units that combine to create larger ones, and it is governed by tacitly understood rules for which combinations are permissible and which are not. The smallest unit of an oral language is a sound, or *phoneme.* Language disorders may, but do not always, include an impairment of the sound system. Even if children do not have difficulty in producing the sounds of their language, they may have difficulty segmenting and recognizing those

sounds as individual units. Awareness of these sound units forms the foundation of phonological awareness, which in turn is the most reliable predictor of early reading success. Phonemes are combined according to language-specific rules to create a slightly larger unit, the *morpheme*, the smallest unit that carries meaning in a language. A morpheme may be either what is commonly considered a word, for example, *jump*, or what is typically thought of as the prefixes or suffixes that shade the meaning of a word. For example, when the past tense suffix *-ed* is added to *jump* (i.e., *jumped*), the meaning of this word is shaded to reflect not only the action but the time frame in which it occurred. Children with a language disorder have particular difficulty acquiring morphemes. Morphemes are then combined in larger units to form the grammar or *syntax* of a language. These units are commonly thought of as simple sentences, or when embedded into one other, a complex sentence. Thus, the form or structure of the language within which we express our ideas to others is completed. The earliest emerging sentences are two-word combinations (e.g., "want cookie"), and late onset of this near-universal stage of language acquisition is often one of the first signs of a language disorder. The potential of a language's form as a powerful means of expression is realized once it intersects with its content or meaning, for without ideas about the world or internal desires to express, the form of language is empty. The linguist Noam Chomsky illustrated this point in his now famous sentence, "Colorless green ideas sleep furiously," exemplifying the notion that phonemes and morphemes may be combined into a sentence that has followed all the rules of a language but expresses nothing. The content component of language, or the ideas we hold about ourselves and other people and things in the world, both draw upon and contribute to the child's conceptual development. Reduced or impoverished vocabulary development is often part of a language disorder.

The third component of the system is use, or *pragmatics*. The central role that language plays in human affairs comes from its communicative function. Arguably, the main purpose of language is for communicating with other people. It is through social interaction with caregivers that the child's genetic potential for language is guided to emerge. It has been suggested that the use of language in social contexts must also be governed by tacit rules, or successful conversational exchanges could not take place. The *cooperative principle* suggests that speakers and listeners have tacitly agreed on a shared goal to successfully exchange information in ways that are maximally relevant to one another. Additionally, contextual variations, such as the social status of those communicating, the circumstances, and the purpose of the communication, are all elements of language use and indeed influence the choice of content and form. For example, a child may use polite forms such as "Please, give me a piece of cake" or indirect requests such as "Gee, that cake looks good" when attempting to get something from a grandparent, but the child may choose a more direct approach when talking to a peer: "Give me a piece of cake too!" A listener may understand a statement to be sarcastic when spoken by a peer but as a reprimand if uttered by someone in a position of authority (e.g., "Do you always dress that way when you come to school?!"). A disruption in the use of language in context is most clearly seen in the language disorders experienced by, but not limited to, children with autism spectrum disorders.

Defining Language Disorders

Based on the previously defined components of the linguistic system, language disorders may then be thought of as a disruption in any one or all of them, making it a heterogeneous category of impairments. It may occur solely in the modality of expression, or it may include difficulty in understanding or comprehending form, content, or the context-sensitive variations described in the previous section. Language disorders are most often developmental in nature and of unknown origins, but they can also be the result of an acquired brain injury. Language disorders may occur in the absence of any other known condition and coexist with otherwise normal cognitive function, in which case it is identified as a specific language impairment (SLI). A language disorder also may accompany other developmental disabilities, such as hearing loss, cognitive impairments, cerebral palsy, autism spectrum disorder, or attention deficit hyperactivity disorder (ADHD). Furthermore, it may accompany the impact of certain negative environmental circumstances such as neglect and abuse. Regardless of whether or not there are associated disorders, a language disorder places the child at risk for negative social consequences because of the primary role that language plays in both social interaction and as a vehicle for the acquisition of knowledge. Both spoken and written forms of language may be affected by a language disorder. Indeed, the very nature of the disorder may change over time with changing contextual demands. Specifically, whereas children

may initially be identified with a spoken language disorder during the preschool years, with treatment, the difficulties in oral communication may resolve. Yet, once these same children start school and are faced with the challenges of learning to read and write, they may once again experience difficulty, the characteristics of the disorder changing with new demands being placed on the language system. Language disorders are chronic disabilities that show remarkable stability throughout the life span.

Prevalence, Stability, and Prognosis

Large-scale studies have been conducted in the United States, Great Britain, and Canada of primarily kindergarten children to establish the prevalence rate of language disorders. In studies of kindergarten children from the general population, the prevalence of language impairment without accompanying speech problems ranges between 6% and 10%. When language problems are combined with speech sound impairments, the rate rises to approximately 12%. In high-risk populations, such as children living in poverty, recent studies have shown a much higher prevalence rate than expected in the general population. Reported stability rates are quite variable for children identified as having a language disorder before they start school (age 4 or 5). It has been reported that approximately 40% of these children will go on to have a good outcome, particularly if only expressive language abilities are affected. These moderate rates of improvement are, however, somewhat illusory when regression to the mean or the impact that false positive identification has on subsequent assessments is taken into account. Furthermore, the prognosis is poorer for children with additional comprehension deficits and below-normal nonverbal cognitive abilities. The bottom line is that for children who are correctly identified as having a language disorder by age 4 or 5, the prognosis for spontaneous recovery is poor. Language disorders are chronic, and longitudinal studies clearly demonstrate that they persist into adolescence and adulthood. Encouragingly, studies investigating the efficacy of language intervention indicate those who have had treatment generally fare better than those who have not, yet much more research is needed to clarify the impact of the amount, nature, and timing of treatment on long-term prognoses.

Given that late onset of single spoken words and protracted periods of vocabulary and grammatical development are hallmarks of language disorders, it would seem that identification of a language disorder during toddlerhood would be ideal so that early intervention could be instituted. Two-year-olds with expressive vocabularies of fewer than 10 words are often considered "late talkers" (about 10%–15% of toddlers); in recent years these children have been widely studied with the hope of determining which of them would "catch up" and which would not. Overall, the current consensus is that the majority of toddlers with early expressive language delay are truly "late bloomers" and they are *not* at risk for further language disorders. In contrast, approximately 40% of late talkers continue to have language difficulties. Although they seem to catch up in their vocabulary development, their continuing difficulties are reflected in difficulties with developing sentence structure. It has been suggested that along with having smaller expressive vocabularies, these children also understand fewer words and use fewer communicative gestures, such as pointing and reaching, as a means of expression. Parent report questionnaires are often the tool used to measure a toddler's emerging language, but unfortunately, they are not sensitive enough to accurately predict persistent language disorder in individual toddlers.

A great deal of attention has been directed to those children who apparently have a language disorder in the absence of any other condition, that is, those children who have SLI. Although there indeed may be a subgroup of children who have a language disorder without any other concomitant condition, the specificity of language disorders in children is debated widely. Many of these children have nonspecific language impairment, that is, concomitant nonverbal cognitive limitations or other associated developmental disorders. Studies have indicated that nearly 50% or more of children identified as having SLI also have associated developmental motor coordination disorder, ADHD, or other emotional or behavioral problems. Thus, in the educational setting, it would not be uncommon for a child with a language disorder to receive intervention by not only a speech language pathologist but also an occupational therapist, an educational resource specialist, or all three.

Causation

Any behavior as multifaceted and complex as language development most certainly must be affected by a variety of biological and environmental factors. Thus, it follows that language disorders can rarely be fully accounted for by one direct and isolated causal

condition. Researchers are just at the frontier of understanding the neurobiology of developmental language disorders, but studies of family history patterns suggest familial inheritance as a factor in at least SLI, particularly expressive-only forms of the disorder. In the past 15 years, the genetics of language disorders, in particular for SLI, has been of great interest, with some researchers proclaiming that a specific gene for grammar (syntax) has been located. However, there is no conclusive evidence that genes can influence the components of language differentially or indeed influence language differently from other cognitive processes. Even though studies with twins demonstrate the importance of genetic influences upon spoken language disorders, researchers are quick to note that just as a shared environment makes the largest contribution to individual differences in typical language development, it has substantial influence in language disorders as well. More research is needed to identify those important contributing environmental variables. Certain sensory and cognitive deficits have long been associated with language disorders in childhood, but association does not mean causation. Hearing loss is a case in point.

It is commonly assumed that a hearing loss causes a spoken language disorder. Yet it has long been known that two children with exactly the same pattern of loss and residual hearing may have quite different spoken language abilities and disabilities. It is not possible to predict language ability based upon an audiogram. In what sense then does the hearing loss cause the language disorder? Similarly, even though language and cognition are intimately interwoven, in cases of cognitive impairment, language abilities may or may not be commensurate with the child's mental age. Indeed, language abilities in advance of cognition characterize children with William's syndrome. So again, one must ask in what sense does cognitive impairment cause language disorder? Although a great deal more research is needed before a satisfactory answer to the question of "What causes childhood language disorders?" is found, it is most likely to reside in the complex interactions of biology and multiple environmental factors.

Language Disorders and Academic Performance

Language is the medium in which classroom instruction takes place. Teacher-directed lessons, small student work groups, worksheets, and textbooks are all language dependent. Moreover, the nature of the language of the classroom is different from that which the preschool child experiences at home. For example, the allocation of speaking turns and topics are teacher controlled as compared with that within the home where conversation is most often child initiated and directed. The language itself is not as closely tied to the immediate context in the classroom, making it necessary to depend more heavily on the meaning inherent in words and sentences themselves. While in the home, the preschool child usually talks about the here and now, but once he or she enters school, much more of the talk is distanced in time and place. Finally, language itself not only becomes the medium by which curriculum is accessed through reading and writing but also becomes a subject of study. While the typically developing child navigates these changes in how language is used when entering school, they present significant challenges for the child with a language disorder. Difficulty following multiple instructions, following classroom "rules" like raising a hand before speaking, or staying on topic can be misinterpreted as behavior problems if educators are not aware of the impact of a language disorder on the functional use of language in the classroom.

Language Disorders and Reading

The contributions of oral language abilities to developing literacy skills have been increasingly recognized in recent years. Indeed one of the most consistent findings in the reading literature is that a child's level of phonological awareness is the best predictor of reading success in the early grades. As described earlier, phonological awareness involves the recognition and manipulation of the smallest units of language. With renewed interest in children's reading comprehension, it also has been recognized that important foundational skills relevant to reading comprehension emerge in their comprehension of oral narratives such as stories well before a child enters school. With the well-established links between oral language abilities and reading, spelling and written composition, it comes as no surprise that longitudinal studies of children with language disorders have revealed long-term difficulty in the language arts areas and even in mathematics. Hugh Catts and colleagues have found that although not all children identified with a language disorder at kindergarten will have reading difficulties, there is a significantly high risk for most of them by the time they reach second grade. Two areas of language

impairment contribute somewhat independently to their reading difficulties. Impaired phonological awareness abilities contribute strongly to difficulties in learning to decode written language, whereas impairment in the semantic and syntactic components of oral language contribute more to difficulties in reading comprehension. Indeed, second-grade children who demonstrate poor reading comprehension but intact phonological awareness have been shown to have an oral language disorder that was not identified in their preschool years. Better reading outcomes can be expected for children whose oral language improved between kindergarten and second grade, but the fact remains that the more severe the language disorder is, the poorer the reading outcome will be. Thus, although phonological awareness instruction has been advocated for children with language disorders, it is with the caveat that it be part of comprehensive oral language remediation. Persistence of academic difficulties has been revealed in longitudinal follow-up studies. Adolescents with histories of a language disorder in kindergarten continue to demonstrate spelling, word recognition, and reading comprehension deficits. When interviewed about their level of literacy, they report that while they read regularly, they do not always understand what they have read and that their written language skills do not meet the expectations of classroom curriculum. Even those whose language impairment appeared to have resolved by kindergarten continue to demonstrate a high rate of literacy problems.

Finally, whereas poor reading outcomes are easily understood given the nature of language disorders, the reported difficulties that children with language disorders have with mathematics are not necessarily as intuitive. Children with language disorder are at risk not only for difficulties with the more obvious areas of "word problems" and when asked to verbally explain their solutions to math problems but also for the rote memory aspects of math needed for number recall and calculation. The relationships between language disorders and success in the academic setting underscore the serious long-term impact of a language disorder, its consequences to both the individual and society.

Language Disorders and Socialization

An account of language disorders in children would be incomplete without a consideration of its impact on social competence. The school setting is highly influential in a child's social development and reciprocally depends upon it. Kindergarten teachers identify social maturity as more important to school readiness than skills like counting and letter recognition. School becomes a primary environment for peer interaction and the development of friendships. Thus, once again, language, as the medium for interpersonal interactions, plays a critical role in the development of social competence, and children with language disorders are disadvantaged.

The social communication difficulties of children with language disorders, while variable, are well documented. These children may demonstrate difficulty initiating interactions with other children and may not be responsive to verbal overtures from peers; they may not be able to gain access to a playgroup or to be integrated into classroom work groups. The general social behavior pattern appears to be one of reticence and withdrawal, which make friendship formation challenging and peer acceptance problematic. As early as the preschool years, other children recognize children with diminished language abilities, and their social interactions are negatively affected. Although not well studied, there are reports that children with language disorders may experience peer victimization and bullying. The precise relationship of language disorders to social competence is a matter of some debate, but the elevated risk for negative social outcomes is not. The limited longitudinal studies that have examined psychosocial and quality of life issues in 20- and 30-year-old adults who have a history of language disorder suggest evolving consequences. The limited data available to date suggest that at least a cohort of mostly males in their 20s, who not only have a history of language disorders but continue to demonstrate diminished language abilities, are satisfied with their lives. It was suggested, however, that when these young men experienced the full weight of adult responsibility, satisfaction ratings might change. Indeed, longitudinal research out of Great Britain revealed that adults in their 30s with a history of severe developmental language disorders indeed had poorer social adaptation and psychosocial adjustment. These adults experienced diminished personal relationships, difficulty maintaining employment, and even difficulty with bullying from coworkers. Although any individual's personal outcome will surely vary, the data are clear on the long-range prognosis for the social competence of individuals with developmental language disorders.

Childhood language disorders are a chronic condition with significant impact upon social, academic, and vocational success. Yet there is evidence, and more is accumulating, that intervention early and tailored to the changing facets of the disorder throughout the school years can improve outcomes for these individuals.

Elizabeth Skarakis-Doyle

See also Communication Disorders; Disabilities; Dyslexia; Learning Disabilities

Further Readings

Bishop, D. V. M. (2004). Specific language impairment: Diagnostic dilemmas. In L. Verhoeven & H. van Balkom (Eds.), *Classification of developmental language disorders: Theoretical issues and clinical implications* (pp. 309–326). Mahwah, NJ: Lawrence Erlbaum.

Brinton, B., & Fujiki, M. (2005). Social competence in children with language impairment: Making connections. *Seminars in Speech and Language*, 26, 151–159.

Catts, H. W. (1993). The relationship between speech-language impairments and reading disabilities. *Journal of Speech and Hearing Research*, 36, 948–958.

Catts, H. W., Fey, M. E., Tomblin, J. B., & Zhang, X. (2002). A longitudinal investigation of reading outcomes in children with language impairment. *Journal of Speech, Language and Hearing Research*, 45, 1142–1157.

Catts, H. W., Fey, M. E., Zhang, X., & Tomblin, J. B. (1999). Language basis of reading disabilities: Evidence form a longitudinal investigation. *Scientific Studies of Reading*, 3, 331–361.

Fazio, B. B. (1999). Arithmetic calculation, short-term memory, and language performance in children with specific language impairment: A 5-year follow-up. *Journal of Speech, Language and Hearing Research*, 42, 420–431.

Johnson, C. J., Beitchman, J. H., Young, A., Escobar, M., Atkinson, L., Wilson, B., et al. (1999). Fourteen-year follow-up of children with and without speech/language impairments: Speech/language stability and outcomes. *Journal of Speech, Language and Hearing Research*, 42, 744–760.

Nation, K., Clarke, P., Marshall, C. M., & Durand, M. (2004). Hidden language impairments in children: Parallels between poor reading comprehension and specific language impairment? *Journal of Speech, Language and Hearing Research*, 47, 199–211.

Naucler, K., & Magnusson, E. (1998). Reading and writing development: Report from an ongoing longitudinal study of language-disordered and normal groups from pre-school to adolescence. *Folia Phoniatrica et Logopaedica*, 50, 272–282.

Records, N. L., Tomblin, J. B., & Freese, P. R. (1992). The quality of life of young adults with histories of specific language impairment. *American Journal of Speech Language Pathology*, 1, 44–53.

Stothard, S. E., Snowling, M. J., Bishop, D. V. M., Chipchase, B. B., & Kaplan, C. A. (1998). Language-impaired preschoolers: A follow-up into adolescence. *Journal of Speech, Language and Hearing Research*, 41, 407–418.

Tomblin, J. B. (2006). A normativist account of language-based learning disability. *Learning Disabilities Research & Practice*, 2, 8–18.

Latinos

See Hispanic Americans

Learned Helplessness

Learned helplessness is a term coined by Martin Seligman to characterize the overgeneralized learning of helpless responses that occur when animals are repeatedly exposed to noxious, uncontrollable, and inescapable situations. What is learned is that their actions will not effect the outcome they desire. In humans, learned helplessness is particularly problematic in education; students may falsely believe that effortful actions, such as studying, will have no effect on performance or learning.

Seligman and colleagues' explanation of learned helplessness focuses on three components: outcome contingency, mediating cognitions, and behavior outcomes. Contingency concerns individuals' perception that outcomes are *contingent* on their behavior. Individuals "learn" helplessness when experiences create the belief that outcomes are not contingent upon their actions. Learned helplessness theory also postulates that certain types of cognitions, called *causal attributions*, mediate this relation between experience and learned helplessness. Finally, learned helplessness theory postulates that the *behavioral* outcomes of perceived helplessness include passivity, quitting, and depression.

The idea of learned helplessness has helped researchers and educators understand why some students repeatedly experience more failures and give up—often before even trying. The serious implications of learned

helplessness include the failure to initiate action, failure to learn, and emotional problems such as depression. There are several ways that teachers can influence students' outlooks on learning, including attribution retraining, encouragement, and focusing on mastery goals.

Research

Seligman and colleagues were the first to manipulate learning deficits by exposing animals to inescapable aversive stimuli. Their phrase *learned helplessness* was used both as a description and a theoretical explanation of the observed learning deficits. In typical learned helplessness laboratory studies, researchers would first expose dogs to a series of inescapable electric shocks. Up to 24 hours later, the researchers put the dogs in the same situation except that they could now avoid or terminate the shocks by simply stepping over a small fence into another part of the training box. If the dog did not terminate or try to escape from the shock, Seligman and colleagues designated it as exhibiting learned helplessness; they believed the dog had learned that nothing it could do would matter.

Researchers extended this animal model of learned helplessness to humans, although the experiments were considerably less aversive. Experiments with humans would present irritating noises instead of electrical shock, but the results were still similar to the results with dogs. In one classic study, researchers exposed participants to a loud noise. During the first set of trials, the participants in the contingent control group could stop the noise with the simple pressing of a button; in contrast, participants in the noncontingent "helpless" group could not stop the noise. In the second set of trials, both groups could move a knob to avoid a signaled noise. Despite the change in contingencies, the participants previously exposed to the uncontrollable noise sat passively and did not attempt escape. In contrast, participants from the response contingent control group and an additional no-noise control group quickly learned to escape the noise by simply moving the knob. Seligman and colleagues concluded that humans, like dogs, develop learned helplessness when exposed to uncontrollable aversive stimuli. Further research on learned helplessness in humans shifted to using unsolvable problems or puzzles as the uncontrollable aversive stimuli. Researchers have subsequently found remarkable individual differences in children's responses to

unavoidable failure on cognitive tasks, which has led to the educationally important question of why some succumb to learned helplessness and others do not.

Attributions and Learned Helplessness

Two children, each of equal ability, are given a puzzle to solve. As long as they are successful, both appear to be motivated and energetic in their approach. But, if this flow of success is interrupted with a set of very difficult or unsolvable tasks, two divergent motivational patterns might be seen. Some children face and recover from failure in a manner that is highly resilient. They appear emotionally unaffected by their experience of failure, simply using failure as an additional source of information. Their performance quickly rebounds once given solvable problems. In contrast, other children appear to crumble in response to perceived or even anticipated failure. These nonresilient children may regress to a lower level of skill than they began at, exhibit negative affect, and conclude that they lack ability. What causes one child to respond resiliently to failure while another gives up? One explanation is attribution theory; differences in susceptibility to learned helplessness are the result of the causal attributions persons make for the aversive situation. An attribution is the causal explanation one makes to explain events. Attribution theory's explanation of learned helplessness asserts that when persons perceive noncontingency, they search for a causal attribution to explain it; this attribution then mediates their response. For example, if failure is attributed to a lack of effort, the response will be resilient. If failure is attributed to lack of ability, then learned helplessness will result.

After considerable research on learned helplessness in humans, Lyn Abramson, Seligman, and John Teasdale reformulated the animal-based model of learned helplessness. This reformulated model, drawn from attribution theory, uses three dimensions of attribution to explain the generalization and impact of perceived failure on performance. The first dimension is *locus on control*, which defines the cause of an outcome as being either internal or external to the individual. For example, the teacher behavior would be an external cause of success or failure, but individual ability would be an internal cause. The second dimension of

attribution theory is *stability*. Some causes of success or failure are assumed to be stable, whereas other causes can change. External causes such as luck are generally believed to be unstable, but internal causes such as ability are often believed to be very stable. The stability dimension is linked to success expectancy. If success is attributed to unstable causes (i.e., luck), then students may believe that their chance of success is unlikely. *Globality* is the third dimension of attribution theory. A student may attribute failure to a global cause ("I'm just not good at math") or to a specific cause ("I only failed because I was too sick to come to class!"). Global attributions assume helplessness will occur in a broad number of situations, thus implying that helpless situations are unavoidable. In contrast, specific attributions confine helplessness to very specific situations, thus implying that success is possible with only small situational modifications.

Attribution theory posits that the causal attributions (internal or external, stable or unstable, global or specific) persons make about aversive situations determine susceptibility to learned helplessness. Motivational deficits associated with learned helplessness include task withdrawal or avoidance. Emotional deficits associated with learned helplessness include sadness, fear, frustration, anxiety, and depression. For example, global and stable attributions generalize learned helplessness across tasks and time. In contrast, specific and unstable attributions limit the generalization of learned helplessness. Stable and global attributions for failure often contribute to helpless behavior, but internal attributions for failure can go either way. For example, effort is an internal attribute, but because it is also unstable (you can apply it, or not), blaming failure on a lack of effort does not tend to lead to helplessness. Interestingly, stable, global, and internal attributions for success increase resiliency to subsequent failure.

Seligman, Alloy, Abramson, Teasdale, and others have developed a learned helplessness theory of clinical depression based on attribution theory. This theory suggests that persons who tend to make stable, internal, and global attributions for negative events are vulnerable to depression. Depressed individuals often believe that they have little control over either negative or positive events in their lives. In sum, the attributional theory of learned helplessness asserts that individual differences in vulnerability to helplessness, and even depression, are the result of individual differences in attributional style.

Current Perspectives on Learned Helplessness

The learned helplessness explanation of performance deficits following perceived failure focuses on beliefs about the contingency of behavior and outcomes; in plainer words, "I have not succeeded because it is not possible for me to succeed . . . therefore I give up." This, however, is too simplistic an explanation of differences in achievement resiliency. Consider the boy who, after studying hard, fails at his first math test of the year, then consequently refuses to study for the next test. Is learned helplessness the best explanation for this motivational change? A significant body of thinking and research has argued that ego threat, not perceived noncontingency, is the real motivator. From this ego threat perspective, the need to protect one's sense of self-worth from displays of low ability can motivate a variety of behaviors, including the task withdrawal described by learned helplessness researchers. Recent research in educational psychology has found ego or self-worth maintenance to be a more fertile approach to achievement motivation and failure resiliency than just learned helplessness. Attribution theory and learned helplessness are integral aspects of current achievement motivation theories, but they are not adequate stand-alone explanations. Another interpretation of helplessness is physiological rather than cognitive. That is, when inescapable shock is presented, there is a decline in the neurotransmitter norepinephrine. This decline has been associated with helplessness behavior.

Learned Helplessness in Students With Learning Problems

Research has found that susceptibility to learned helplessness varies considerably from individual to individual. Learned helplessness is related to a variety of individual characteristics: anxiety, depression, low self-efficacy, low mastery and high performance orientations, low self-worth, extrinsic motivational orientation, and need for structure, to name a few. Students with learning problems, however, may be particularly vulnerable for learned helplessness. Students who repeatedly encounter academic difficulties come to expect failure, which generalizes to low self-efficacy, which inhibits effort, which guarantees further academic failure, which, by reinforcing failure expectations and low self-efficacy, returns the student to the beginning of this vicious cycle. Even when success is

experienced, these students may attribute it to external causes, for example, believing that they succeeded at a math problem because it was easy, because they received help, or because they were simply lucky. However, not all students with learning problems display helplessness; some are quite resilient. The frequency of failure may be an important factor: Students who fail in multiple domains may be more susceptible to developing learned helplessness. In addition, reading difficulties are especially important, as reading skills determine success in many other domains of competence.

What Can Educators Do?

Teachers can provide students with feedback on the attributions they make about the causes of their success and failure, and they can model healthy attributions. Progress toward goals and success should be attributed to internal attributes such as effort and ability, but failure should be attributed to internal unstable attributes such as effort. In addition, teachers should provide moderate difficulty tasks that students are able to accomplish with effort. Success on tasks that are too easy does not produce healthy ability or effort attributions, and failure on tasks that are too difficult usually results in unhealthy, low ability attributions.

Teachers can help students by explaining the role that attributions play in the learning process. For example, students can be encouraged to focus on controllable factors such as effort and strategies, rather than uncontrollable factors such luck or mood. Students can be taught to consider other causes of success and failure. Students experience difficulties in class for several reasons. For example, these reasons may include not using appropriate learning strategies, or lacking prior knowledge. These factors are often controllable, and students should recognize that many failures may be attributable to these changeable factors. Research shows that attribution retraining is an effective intervention for learned helplessness. By increasing students' awareness of the attributions they make, their self-esteem, self-efficacy, and learning can be improved, and their anxiety and frustration can be reduced. Another approach is to encourage children to focus on mastery goals, or learning for its own sake. Mastery goals increase actual competence, which in turn increases self-worth and achievement motivation. Focusing on grades, the approval of others, and competition are performance goals. Performance goals work fine as long as the student is winning; however, failure at a performance goal can produce ego-protective maneuvers that interfere with learning. Over time, performance goals make learners susceptible to learned helplessness.

Jennie K. Gill and Joan M. Martin

See also Failure, Effects of; Learning; Motivation; Self-Efficacy; Self-Esteem

Further Readings

Abramson, L. Y., Seligman, M. E. P., & Teasdale, J. D. (1978). Learned helplessness in humans: Critique and reformulation. *Journal of Abnormal Psychology*, *87*, 49–74.

Alloy, L. B., & Abramson, L. V. (1982). Learned helplessness, depression and the illusion of control. *Journal of Personality and Social Psychology*, *42*, 1114–1126.

Buchannan. B. M., & Seligman, M. E. P. (Eds.). (1995). *Explanatory style.* Mahwah, NJ: Lawrence Erlbaum.

Maier, S. F., & Seligman, M. E. P. (1976). Learned helplessness: Theory and evidence. *Journal of Experimental Psychology: General*, *105*, 3–46.

Seligman, M. E. P. (1975). *Helplessness: On depression, development, and death.* San Francisco: Freeman.

Learning

Human learning has been the focus of organized study for many decades, and the results of this work have become ever more important as societies intervene on so many levels to promote and influence learning. Today there is no one single way to define *learning*. Rather, what is found is a range of explanations, each of which provides an important frame of reference for thinking about learning as a human endeavor. Generally speaking, over the past 60 years, three major conceptual frameworks have emerged, and these three will be the focus of this entry.

The first of these frameworks looks at learning in terms of observable behavior. In simple terms, learning is defined as any relatively permanent change in behavior that is not the result of normal growth or maturation. There is no limit to the range of behaviors that might be considered or the contexts in which they occur. When people drive automobiles or operate machinery, perform school tasks such as writing and calculating, or engage in social activities with others, generally their behaviors are fairly complex behaviors that have been acquired over time and with much

practice. People who study learning from a behavioral perspective want to know how these complex behaviors are acquired and how they change over time.

The second framework, which began to appear in the late 1960s and early 1970s, deliberately moved away from behavioral explanations, focusing instead on the information-processing activities that occur in the human brain. This movement, known as the *cognitive revolution* in human learning, developed largely around questions about memory and meaning. For example, when a person listens to people talk—or reads from text—how does he or she process and store what was heard or seen? How does the person represent information in memory for later use, and how does he or she gain access to the large amount of stored data? These and other related questions have dominated the study of learning for many decades, and they continue to be prominent in researchers' thinking about learning. Thus, viewing learning as a cognitive activity, it can be defined as the acquisition of knowledge and the ability to use knowledge to solve problems.

A third framework for investigating human learning began to be noticed during the early 1990s. In contrast to the cognitive point of view, in which learning is defined in terms of an individual's computation of information, this framework focuses more on how people work and learn in cultural settings. Here learning is defined not as the acquisition of knowledge but as participation in meaningful social practices. Examples of cultural practices naturally include a broad range of activities, such as child rearing, office work, professional endeavors, trades, hobbies and the like. As people participate in social practices, they develop roles relevant to their particular type of participation, and as these roles develop, people acquire identities as legitimate practitioners. One important distinction between this framework and the cognitive viewpoint is that in this framework, learning from a social perspective is never separated from doing.

When looking across this history of progress in learning research, there is a natural temptation to simply focus on the latest prominent explanation that may be enjoying most attention. Certainly today there is very little discussion of behaviorism as a viable framework for understanding learning. In fact, the intention and purpose of the cognitive revolution was not simply to modify behaviorism but to replace it altogether.

It can be argued, however, that each of the three frameworks outlined here provides an important window on learning and allows us to see human learning as the rich and multifaceted phenomenon it really is.

Learning as Behavior Change

Learning from a behavioral perspective involves examining the relationship between what animals or humans do (i.e., their responses) and what the environment does (i.e., various stimulus events). When behavior changes systematically in relation to environmental stimuli, behaviorists refer to such learning in terms of a *conditioning* process. A modest complexity in this is that two types of conditioning are possible, and they are based on different learning principles.

Classical Conditioning

Based on the work of Ivan Pavlov, the classical or Pavlovian model is rooted in simple reflexive behavior. The basis for this model lies in the range of relatively permanent and unlearned reflexes that nearly all members of a species possess. Examples include our automatic reactions to hot surfaces, food smells, and a whole range of stimuli that may cause fear, anxiety, flight, or a sense of well-being. These automatic reflexes are composed of two elements: an *unconditional stimulus* and an *unconditional response*. Again, no learning is required for these responses to take place. Where the learning, or conditioning, comes in is when another neutral stimulus is introduced in just the right way when these automatic reflexes are occurring.

The classic experiments from Pavlov's laboratory involved his work with the digestive process in dogs. As the story goes, Pavlov noticed that not only would dogs salivate in the presence of food (this is the unconditional stimulus-unconditional response reflex), they also would initiate the salivary secretions even before the food arrived. This curious phenomenon set Pavlov on a course of experiments in which he discovered that various neutral stimuli, such as the sound of the attendant carrying the food or the sight of the food bowl, were enough to induce the dogs to salivate. This learning process simply required that the neutral stimuli occur in some close relationship with the food itself. Essentially then, the events just prior to actual feeding became conditional stimuli causing the conditional response of salivary secretions.

Today, Pavlovian-type conditioning is seldom discussed in the context of learning. This is unfortunate, in part because such conditioning is probably at the

root of much emotional learning that occurs in schools and in daily life. Think about the fears and anxiety that adults and children experience in medical offices and hospitals or in school-related environments, such as classrooms and school busses. If, in these environments a person directly experiences pain, failure, or hostility, the emotional response is not simply to the needle that caused pain, the test result that caused shame, or the fight that induced fear, but it is spread to surrounding cues such as the smells of the hospital or uniformed personnel, the classroom in which poor math performance occurred, or the vehicle in which the fight or hostility occurred. This type of emotional learning can be underappreciated as a very explainable phenomenon, tied directly to the classical conditioning model.

Operant Conditioning

The better known and broadly applied behavioral model is connected with the pioneering work of E. L. Thorndike and B. F. Skinner. This approach to learning is based on the simple proposition that behavior is a function of its consequences. The earliest expression of this principle is Thorndike's Law of Effect, which observes that when any response is followed by a satisfactory state of affairs, it tends to be repeated. Similarly, when a behavior is followed by an unsatisfactory or annoying result, it tends to reduce in frequency and strength.

Building on the Law of Effect, Skinner and others created an elaborate set of ideas that explained how to increase behaviors of interest, how to teach totally new behaviors, how to maintain learned behavior, and how to reduce unwanted behavior. All of these explanations rested on understanding the relationship between three elements of the puzzle: the behavior of interest, the antecedent conditions in which the behavior of interest was occurring, and the consequences following behavior. The relationships among the three elements are sometimes referred to as the *contingencies of reinforcement.*

Building on the central idea that behavior is a function of its consequences, the next step is to examine four types of possible consequences and their effects. Two of these, *positive reinforcement* and *negative reinforcement*, tend to strengthen or increase behaviors because they result in a positive state of affairs for the subject. *Punishment* can be used to suppress behavior, but the most effective long-term solution to

reduce unwanted behavior is simply to deny reinforcement. This procedure is referred to as *extinction.* The most effective combination of these tools would be to use positive reinforcement to strengthen desired aspects of a subject's behavior, while simultaneously using extinction to reduce less desired or competing behaviors. An important point to note is that punishment can be effective in the short run, but it is never a desirable long-term solution because of the negative side effects that can result, such as hostility and fear. Certainly, one of Skinner's most important contributions was his advocacy of positive consequences in the control of learning.

Learning episodes that involve operant conditioning are usually not just about increasing or reducing behaviors but also about producing responses that are appropriate to particular situations. This is where the concept of *stimulus control* comes into play. People respond predictably to particular stimulus situations (e.g., traffic signals, classrooms, restaurants) because they have been systematically reinforced for appropriate behavior in those settings. Following such selective reinforcement, the stimulus situation itself comes to control behavior, as when a red light shows at a traffic intersection or people politely stand in line at a ticket counter.

Learning is also a matter of acquiring completely new behaviors that are not presently in existence. Language-delayed children, for example, may be unable to make certain sounds that need to be part of their speech repertoire. If the children could correctly produce the sounds even occasionally, the strategy would be to simply reinforce the correct behavior. However, if the sounds are never produced, the strategy is to use a *shaping* procedure in which the closest approximations are reinforced. Over time, with skilled use of reinforcement, the approximations become closer and closer to the target. This is the same strategy used in animal training, where the object is to teach behaviors the animal is capable of but are not part of their natural repertoire. Again, the strategy is to selectively reinforce closer and closer approximations until the target is reached.

Finally, one important issue in behavioral learning is the maintenance of behavior at desired strength and frequency levels. This is where *reinforcement schedules* come into play. This is a complex area in itself, but the main idea is that constant reinforcement never produces the strongest and most extinction-resistant behavior. Rather, once a behavior is learned and can

reliably occur in a particular stimulus environment, the reinforcement schedule needs to be adjusted so that reinforcement occurs only intermittently. A good example of this is the timing of payoffs in the gaming industry, particularly with slot machines. Patrons often spend long hours on these machines because the schedule is cleverly arranged to keep gamers looking for the reinforcement. Teachers and parents too quickly find that if children are constantly reinforced, the behavior in question will quickly disappear once the reinforcement is removed. This would not be the case had the behavior been reinforced on a leaner and more unpredictable schedule.

The entire operant conditioning framework provides a powerful set of tools for anyone who desires to intervene in his or her own behavior or to work toward change in others. To recap, the main idea is that behavior does depend on its consequences. Use of the operant conditioning principles is not difficult, but in the absence of a basic understanding, it is easy to make mistakes.

Learning as Information Processing

During the long period in which behavioral frameworks dominated learning research, the activity of the human brain was somewhat of a mystery. However, as described earlier, the many questions about what goes on inside the "black box" began to be revealed as investigators employed a range of methodologies to tease out some of the basics of how humans process information and how the memory system worked. This movement began in the late 1950s and was experiencing great momentum by the 1970s. In a nutshell, the entire conversation about learning rapidly shifted from questions about behavior to questions about knowledge acquisition and the construction of meaning.

This section outlines the basic themes and highlights the main contributions of the cognitive orientation to learning, as follows:

1. It is important to recognize the active nature of human information processing and to understand some of the capacities and constraints of the memory system.

2. Because knowledge acquisition is such a central theme in cognitive learning, it is important to know something about the structure and organization of memory. If learners are constantly building knowledge, what do their mental representations look like?

3. Learners necessarily exercise a great deal of control over their own learning, so it is important to consider that people have much to learn about their own learning.

4. Learning is not simply a linear process in which people continue to add to their knowledge stores; rather, learning is a cyclical process of returning again and again to existing concepts that must be refined and oftentimes corrected for misconceptions.

Active Processing and Capacity Issues

The original memory models posited several types of memory stores. For the purposes of this entry, it is sufficient to distinguish between a short-term "working memory" and a long-term, or permanent, memory. Short-term memory is the conscious workspace where new information is briefly held in an active state so a person can either use it for quick responses or process it further into a more permanent form. One difficulty with short-term memory is that it is plagued by limited space and also limited time duration. The general rule is that short-term memory can hold about seven, plus or minus two, bits of information, and without constant mental rehearsal, the duration is slightly less than 30 seconds. Once a person processes information into long-term memory, these constraints disappear; however, other issues arise, such as how one goes about finding information that was previously stored.

These complexities of memory simply illustrate the need for strategy and skill in using the system. With short-term memory, learners can keep information "alive" longer with rehearsal strategies or by using external stores, and the capacity problem can be partially solved by recoding small bits of information into larger chunks based on prior experience. Furthermore, when people process new information for more permanent storage, they are typically connecting it with knowledge stored at a prior time. The quality of these connections makes a big difference in whether the new information is retained and whether it can be retrieved for later use.

Knowledge Representation

Because there is such a vast amount of information stored in human memory, the issues of structure and organization are important. For the sake of space, knowledge representations should be efficient and

avoid redundancy. For example, when a person stores information about forms of literature, there is no need to store separately for each story, poem, or play, as literature has a basic structure that is relatively constant across thousands of examples. Rather, with experience the person develops basic schemas for how plays, poems, and stories generally go. These schemas may change somewhat over time, but the important thing is that when the person encounters new examples of literature, he or she knows a lot in advance about what to expect.

A second benefit of well-formed knowledge is that it gives learners the power to generate their own knowledge. When students in any discipline master the basic underlying structure of their disciplines, they can then generate inferences and further details that were never directly taught, but that are reasonable extensions of the basic knowledge base.

Several different types of knowledge representations are represented in the literature on learning. Everyday affairs such as eating at restaurants, going to doctors and dentists, participating in a religion, and observing holidays have underlying structures that are usually temporal in nature. These structures are typically referred to as *scripts*. People depend on these underlying scripts to give order and predictability to regularly encountered experiences.

By contrast, the facts and concepts learned in schools tend to be represented in terms of *propositional networks*. Researchers can easily demonstrate the "networked" nature of basic ideas in any discipline simply by naming one example of a concept—the word *disease*, for example—and then asking people to generate related ideas that come to mind.

Probably the most powerful general schema that people anywhere possess is the knowledge of how stories are organized. Because so much of human experience is shared via stories, it is important for all members of any society to be able to understand the elements of story narratives and to be able to generate them for communication. It could be argued that the ability humans have to organize and communicate human experience at a level above mundane details is their greatest achievement.

Executive Control

It is clear that learners control their own cognitive activity, and for this they need two things. One necessity is *awareness* of how they are doing. For example, when people read text material, comprehension is key to their success. Most people have experienced completing a reading passage only to recognize that they have no clue as to what they have just read. This feeling of comprehension, or lack thereof, is the first requirement for being able to take corrective action. Successful readers and learners tend to maintain such awareness to a greater extent than less successful readers.

The second necessity is to acquire and use *strategies* for learning, comprehension, and memory retrieval. This is the area in which learners can vary dramatically in their approach to learning. The basic idea is that people can always "do something" to actively process information, and over time researchers have identified many powerful domain-specific strategies that learners either acquire on their own or can be taught. The bottom line is that nothing substitutes for taking charge of one's own learning.

Learning Over Time

The final point to make in this section on learning from a cognitive perspective is that knowledge building is a never-ending process. Even the straightforward scripts people develop in areas such as their participation in medical services undergo substantial change as practices are modified based on technology and medical science. Whenever one's experience changes in any domain, the underlying schemes or frames used to make sense of those practices must change as well. In other words, the basic story lines that give meaning to people's lives are constantly under construction. In the final analysis, this is what the cognitive world is really about; it's about a constant search for meaning.

Learning as Social Participation

Until the late 1980s the most significant contributions to research on learning occurred in laboratory settings, and the largest application contexts were in the public schools at all levels. As seen in the discussion thus far, the emphasis on learning—whether with respect to observable behavior or less visible cognitive activity—has been on individual learners performing mostly solo. This emphasis on individual performance has served people well in some respects; after all, success in school is largely a solo affair, particularly when it is explicitly reinforced that way with the continuing emphasis on grades as the primary indicator of learning.

What changed during the 1990s is that new research was conducted not in psychology laboratories but by social anthropologists in everyday cultural practices. Examples of the contexts studied included craft cultures such as meat cutters and tailors, professional settings such as naval navigation and insurance claims processing, and other practices such as embodied in Weight Watchers clubs and Alcohol Anonymous meetings. Not surprisingly, the story about learning that emerged from these kinds of studies was framed in an entirely different way, and the details provide a window onto learning that enriches researchers' understanding substantially. One key premise is that humans are social beings, and learning is inherently a social process. This basic framing idea leads to a range of new propositions about learning, which are explained briefly in this concluding section.

The Centrality of Participation

One prevailing mind-set about learning that has been carried forward through the behavioral and cognitive models is that learning is about the acquisition of generic skills and knowledge, with little attention to the contexts that are relevant to these outcomes. In communities where social practices are ongoing, the emphasis shifts to one's role in the practice and the participation structures that afford access to the practice. With this shift, learning is never separated from doing, and in fact learning is defined in terms of one's trajectory as a legitimate practitioner. In this context, beginners may enter a complex practice by initially assuming peripheral roles, but even at the periphery, the novice participant has access to the full social apparatus of the practice—such as practice-related conversations, the tools and contexts of the practice, and members who are more advanced in their roles. Over time. novices move to increasingly central roles, and with this movement, they begin to acquire an identity as a legitimate practitioner.

Identity

If the purpose of learning as cognitive activity is to change people based on what they know, the purpose of social participation is to change people in terms of who they are. The concept of identity is difficult to consider in formal learning contexts, such as schools, because the subject matter of schools (i.e., the traditional disciplines) generally does not allow access to the underlying practice domains referenced by the disciplines themselves. Thus, we could ask how one "becomes" a scientist when the study of science is more about being a student of science than it is about the legitimate practice of science. It is clear, however, that going to school is a form of practice, and it is likely that students, over time, develop identities as people engaged in the practice of school going.

Support for Learning

Learning in communities of practice is generally a context-rich affair. This is in stark contrast to learning in classrooms, where knowledge to be learned is generally presented, often for efficiency, as the bare facts and concepts carefully separated from the messy contexts in which such knowledge was discovered and might be applied. In communities of practice, there is by definition access to all the tools of the practice, all the social structuring, all the stories, and all the physical aspects of the setting that provide clues to how one successfully participates. In these settings the very cognitive activity, such as math calculations, may look entirely different from the same operations as they would occur in school lessons.

Emphasis on Situated Cognition

One immediate consequence of the social perspectives outlined here has been the emergence of a significant debate about the continued framing of learning as a process whereby individual learners acquire knowledge and skill through the addition of small, incremental pieces. The question for formal education concerns how much of the social practice world to let into traditional learning environments. Although this debate is ongoing, it does seem clear that researchers, theorists, and educators can no longer afford to ignore the fundamentally situated nature of learning.

Terry M. Wildman

See also Educational Technology; Effective Teaching, Characteristics of; Emotion and Memory; Episodic Memory; Explicit Memory

Further Readings

Bruner, J. (1990). *Acts of meaning.* Cambridge, MA: Harvard University Press.
Chance, P. (1994). *Learning and behavior.* Pacific Grove, CA: Brooks/Cole.

Gardner, H. (1985). *The mind's new science: A history of the cognitive revolution.* New York: Basic Books.

National Research Council. (1999). *How people learn: Brain, mind, experience, and school.* Washington, DC: National Academy Press.

Skinner, B. F. (1968). *The technology of teaching.* New York: Appleton-Century-Crofts.

Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity.* Cambridge, UK: Cambridge University Press.

LEARNING COMMUNITIES

Learning communities are a family of programs, typically implemented at the collegiate level, that intentionally create community to accomplish specific learning objectives or goals. They are characterized by intentionally crafted spaces that bring groups of learners together to achieve shared goals around a central theme, increase interaction and collaboration with peers and faculty both in and out of the classroom, and use active learning techniques to focus on multi- or interdisciplinary concerns. Studies evaluating learning communities' effectiveness suggest they promote students' attainment of a wide range of academic and social outcomes important to graduates facing new social, political, and economic challenges.

History of Learning Communities

In their review of student learning communities, Oscar Lenning and Larry Ebbers noted that bringing individuals together for the purpose of creating more powerful environments for learning could be traced back to the first century C.E. However, Barbara Leigh Smith and colleagues have noted that the American pragmatist John Dewey provided the foundation upon which today's learning communities rest. In *Experience and Education*, Dewey called for *progressive education.* This focused on experiential (as opposed to vicarious) learning that situated knowing within the experience of the learner and was concerned with not only the transmission of culture but the development of skills needed for future knowledge production and problem solving. Importantly, Dewey noted that the process of progressive education depended on contributions from the student, his or her peers, and the educator.

In *Experimental College*, educator Alexander Meiklejohn described the University of Wisconsin's Experimental College, a residentially based learning community, or living-learning program, that opened in 1927. Students spent their first year in the Athens-America Curriculum, comparing and contrasting the development of each society's political and social development. In their second year, students pursued Sophomore Regional Studies, an in-depth sociological and anthropological examination of a given locale. Five years after it opened, the Experimental College closed due to funding concerns and controversy about its methods. However, several of the college's original design components remain common in today's learning communities, including (a) multi- or interdisciplinary inquiry focused on a particular theme or problem; (b) the blending of social and academic roles for students, their peers, and faculty, including enhanced opportunities for formal and informal contact through shared community space; (c) college-sponsored co-curricular activities and student organizations; and (d) faculty training in teaching methods akin to today's integrative, learning-centered pedagogies.

In the years that followed, institutions interested in developing learning communities adapted all or part of the philosophical and practical advice of Dewey and Meiklejohn to their unique needs. Several notable examples exist, including Tussman's work at the University of California-Berkeley, patterned on Wisconsin's Experimental College; Evergreen State College (Olympia, Washington), with its emphasis on interdisciplinarity, team-teaching, and collaborative, problem-focused learning; and curricular reforms at LaGuardia Community College (New York City) and Stony Brook University. The successes and failures of these programs informed subsequent innovation, and the range of program types now classified as learning communities has grown.

Types of Learning Communities

Although no definitive typology of learning communities has been developed, most programs can be sorted into one of four categories.

1. Fully team-taught programs typically represent the form of learning community with the highest degree of curricular integration, with a small group of faculty working together to design courses around an interdisciplinary theme. These programs, which are also referred to as coordinated studies programs, may represent a student's entire academic load for a semester or year.

2. Paired or clustered course programs, or block scheduling, join two or more discrete but complementary courses. As such, these programs may or may not be organized under a unifying theme or taught in an interdisciplinary manner. Combining an English composition or orientation/first-year transition course with a course from another discipline is a common example, as is joining related natural, physical, or social science classes.

3. Cohort programs are designed to create smaller learning units within large, preexisting courses. A large chemistry class, for example, might be broken into groups of 15 students (a cohort) who attend a large lecture with their peers but then meet as a cohort in a special recitation or tutorial section. In the federated learning community model, cohort members are accompanied to class by a faculty "master learner," who then teaches the recitation section. Cohorts of students participating in freshman interest groups (FIGs) take a group of courses together, which is often a mix of larger lecture, smaller discussion, and orientation/first-year transition courses united under a central theme.

4. Living-learning programs seek to combine one of the three previous learning community models and the student's residential experience to maximize student and faculty contact and provide greater opportunities for formal and informal interaction around the program's contents. Students participating in a given program are assigned to all or part of a particular residence hall, often with faculty offices, academic support resources, classrooms, and spaces for group study or tutoring located directly on their floor or in their building. Residence hall common spaces are also used for programmatic purposes, such as co-curricular activities and social events.

Importantly, each of these program types may also include out-of-class learning experiences, such as service-learning opportunities, community-based problem-solving pursuits, internships or co-ops, team development activities, or relevant cultural events. In doing so, learning communities purport to support students' collegiate success, most notably academic achievement, social growth, and persistence.

Student Outcomes Associated With Learning Communities

Unfortunately, little quantitative research about the effectiveness of learning communities exists when one considers the number of interventions that are being employed nationwide. Patricia Cross has noted that much of what is reported as causal evidence of the effectiveness of learning communities is actually research showing correlations between student outcomes and experiences that are typical of intentionally designed programs, limiting causality claims. Such research also typically relies upon students' self-reported gains as opposed to more objective measures. Making generalizable statements about the effectiveness of learning communities is further complicated because the majority of research that has been conducted is limited to single-program evaluations. Despite these methodological problems, two classes of positive outcomes associated with learning community participation are increased academic achievement and increased student persistence through students' enhanced integration with the social and academic systems of the institution.

Academic Achievement

Participation in learning communities has been positively linked to the attainment of a wide range of academic outcomes. In reviewing a collection of dominantly single-program studies, Lenning and Ebbers noted gains in students' grade point averages, declines in the number of students on academic probation, growth in skills needed for academic success, and enhanced cognitive complexity. These findings have been corroborated by multi-institutional studies. Researchers have documented higher grade point averages among learning community participants in comparison to nonparticipating peers. Using data from the National Survey of Student Engagement, Chun-Mei Zhao and George Kuh identified a wide range of other desirable intellectual or academic outcomes for learning community participants, including more academic effort, more frequent engagement in active and collaborative learning, and more frequent enrollment in courses that demanded the use of critical thinking skills.

Persistence

A number of studies have suggested that learning communities contribute to student persistence. Most attribute this effect to students' enhanced involvement with the institution's academic and social systems, which Vincent Tinto's model of student attrition argues plays a dominant role in persistence behaviors.

More frequent faculty contact, a hallmark of academic integration, often has been linked to learning community participation. Beyond mere frequency of contact, Tinto noted that student-faculty interactions may be qualitatively improved: Students who participated in learning communities reported more favorable perceptions of faculty. Learning communities also allow peers to contribute to academic integration. Tinto and colleagues found learning community students reported more frequent participation in study groups, a finding corroborated in two of the cases in Jean MacGregor's study, and greater levels of peer accountability for academic issues.

Research on learning communities suggest their contribution to social integration is due, in large part, to the creation of environments through which students gain a greater sense of social support. Tinto noted that learning community members reported higher levels of peer interaction than their peers who were not learning community members, and his work has found that students participating in learning communities were able to build peer networks more readily. Additionally, participation in learning communities appears to contribute positively to students' general sense of satisfaction with their experience on campus, and with peers, faculty, and administration.

Living-learning programs may contribute distinctively to academic and social integration because of their co-location of student residence and features characteristics of nonresidential learning communities. Karen Inkelas and colleagues have noted that on many important indicators, living-learning participants reported more positive experiences and outcomes than their non-learning community peers, including smoother social and academic transitions to college, more frequent academic and sociocultural discussions with peers, greater rates of faculty mentorship out of class, and stronger senses of civic engagement and empowerment. Gains on general education outcomes have been associated with participation in living-learning programs, but most researchers have noted that these indirect effects are mediated through enhanced student involvement with their environments.

Core Practices in Developing Learning Communities

Barbara Leigh Smith, Jean MacGregor, Roberta S. Matthews, and Faith Gabelnick are among the most prolific, contemporary proponents of the learning community movement. Their recent work has suggested that the positive outcomes associated with learning communities, like those outlined in the previous section, are dependent on five core practices. Based on theoretical propositions and empirical findings from the literature surrounding student development, cognition, and learning, each practice provides guidance for program development and implementation.

Smith and colleagues' first core practice is *community*, and it is evidenced at the student, faculty, and institutional levels. For students, community is experienced both socially and academically. First, students recognize community as a sense of inclusion, which connotes membership or belonging to the institution's academic and social realms. Indeed, it has been noted that learning communities can provide otherwise isolated students a ready-made location for forming academically and socially supportive peer relationships, both in and out of the classroom. Community is also experienced epistemologically. Cross has noted that, increasingly, higher education is acknowledging a shifting understanding of the nature of learning that recognizes its nature as a social process. Participation in a learning community demands that certain skills and attitudes consistent with this new view of knowledge be fostered. As such, Tinto noted that learning community participants recognize that they have a role in facilitating not only their own learning but also the learning of their peers.

Just as students benefit from the "community" inherent in learning communities, so too may faculty. Learning community scholars have noted that learning communities can serve as communities of practice for faculty, providing opportunities for faculty development and a location for meeting social or affiliational needs. Finally, Smith and colleagues noted that learning communities may have as their focus a mission of service to the institution's external community, extending their benefits well beyond the campus gates through service learning, community-based problem solving, or action research.

Diversity, or a commitment to multiculturalism and social justice, is the second core practice identified by Smith and colleagues. They have noted that diversity may be promoted and supported in learning communities. First, the authors noted the growth in learning communities focused on serving a group of participants traditionally marginalized in the larger community by virtue of race or ethnicity, gender, sexual orientation, or other social identities. They also

suggested that learning communities focused on critical or postmodern analysis of social concerns provide participants of all backgrounds opportunities to be exposed to a curriculum that is more inclusive than that taught elsewhere on campus. Finally, both Smith and colleagues and Cross have noted that pedagogies employed to develop in-class and out-of-class learning experiences may themselves be more inclusive, welcoming those who have traditionally been isolated from the educational dialogue into a community of learners, including women and ethnic minorities.

Smith and colleagues' third core practice, *integration*, calls not only for the integration of diverse curricula to assist students in formulating a view of knowledge as interdisciplinary, but also for the use of pedagogical techniques that empower participants to integrate knowledge on their own accord. Typically, curricular integration takes places through the examination of a central problem through multiple disciplinary lenses. This practice promotes epistemic growth, in that it challenges participants to redefine how they come to know and solve problems. Integrative pedagogies typical to learning communities, such as collaborative learning, study abroad, and service-learning, among others, support student growth by providing an expanded repertoire of methods to explore real-world concerns.

The use of *active learning* is Smith and colleagues' fourth core practice. Acknowledging the benefits accrued through the use of active learning techniques (i.e., undergraduate research, service-learning, fieldwork, and collaborative teaching and learning practices), Smith and colleagues join Lenning and Ebbers in arguing that well-crafted experiences should appear consistently across learning communities' designs. Smith and colleagues contended that when students and instructors actively explore substantial, meaningful questions in environments built upon the core practices, the roles and capacities of those students and instructors are transformed. Learning, rather than instruction, is ascendant; students and instructors share responsibility for the enterprise of learning; and students' capacity to generate knowledge in collaboration with their peers is enhanced, as opposed to their being passive recipients of knowledge handed down from instructors. Learning communities also hold the promise of extending the temporal boundaries of when active learning occurs. Tinto noted that learning community participants' increased social interactions with each other provided opportunities for learning at virtually any time.

Smith and colleagues' fifth and final core practice is *reflection and assessment.* Itself a form of active learning, ongoing reflection and assessment contributes to students' meaning making and is a way to build community among participants. Smith et al. argue that the task of individual reflection should include a thoughtful analysis not only of the content of what has been learned but also of the processes of thinking and learning. Group reflection within a community can serve as an important mechanism for strengthening the process of learning by reinforcing those connections and relationships critical for collaboration. Smith and colleagues also identified assessment as a critical opportunity for enhancing learning. Formative assessment can be particularly beneficial to learning community members, providing regular opportunities for a learner and her or his instructor to discuss what there is to be learned from an experience, where a given learner's strengths lie, and where further growth can occur.

Tasks and Roles of Faculty and Administrators

Learning communities represent a conceptualization of the learning enterprise that fundamentally differs from the traditional model often employed in undergraduate education, varying in both process and product. Didactic techniques are still the norm on most college campuses, at least in introductory and other lower-level courses, and Paulo Freire's "banking model" of education, in which students are the passive recipients of professors' knowledge, frequently remains the underlying instructional pedagogy. Learning community scholars have noted that learning communities have the potential to transform such practices, but doing so will require the redefinition of roles for faculty and campus administrators.

As faculty members begin working with learning communities, it is often difficult to change from the traditional ways of "being" and "doing" in the academy. The greatest transformation may be pedagogical, as the faculty member moves from an instruction-centered paradigm to one that is learning centered and redefines the nature of power in the learning relationship. It is also important to intentionally design integrative curricula that highlight the analytic strengths (and weaknesses) of multiple, and possibly divergent, disciplines and employ active learning pedagogies. Importantly, this reform is often done in concert with

colleagues from across the institution, requiring a level of collaboration that goes beyond what may be typically experienced in faculty life. Finally, faculty involved in learning communities also must sharpen their skills in working with students: Because of increased frequency and quality of interaction and a reduction of the power difference between teacher and learner, faculty—especially those residential learning communities—often find themselves assisting students with social, personal, and vocational concerns.

Learning communities call for student affairs and academic affairs administrators to collaborate in ways that, according to Charles Schroeder and his colleagues, are all too often unfamiliar because of historical divisions. Implementing learning communities requires both areas of the university to acknowledge that their work is united under a superordinate goal: Enhancing undergraduate education. As such, delineating which responsibilities will be met by academic administrators, which will be met by student affairs administrators, and which must be shared creates space for cultural exploration and an eventual understanding of culture's influence on learning communities' success.

For their part, academic affairs administrators must wrestle with difficult questions related to the implementation and ongoing management of learning communities. If release time is needed, for example, how will it be financially supported, who will authorize it, and how will faculty loads be re-balanced? How will learning community teaching or service be evaluated when faculty seek tenure or promotion? Who will handle programs' administrative burdens, especially shepherding new courses through curriculum committees, ensuring that sections of a given course are scheduled in the appropriate sequence and numbers, and handling linked or cohorted class registrations? Although none of these challenges is individually insurmountable, together they represent a burden that must eventually be managed.

Developing and implementing learning communities often require the efforts of student affairs administrators and may include providing particular services associated with student affairs units, such as career planning and placement, service-learning and internships, study abroad, or residential education. Especially in the case of living-learning programs, university housing provides not only the facilities needed for successful programs, but also the support required to handle housing assignments, deliver a broader

residence education program through professional and paraprofessional staff, address roommate concerns, and manage or maintain ancillary space for programmatic needs. Student affairs staff members may also provide needed consultative services for faculty, especially on techniques for building community, promoting leadership development, and shaping positive student behaviors.

Finally, Anna Goodsell Love has noted that in learning communities where both academic and student affairs divisions are partners, there are many tasks that require collaboration. First, both academic and student affairs administrators must identify, and cultivate the continuing support of, key institutional stakeholders (typically particular leaders) who are champions of the learning community movement. Second, programs' long-term stability must be addressed. Sustainability is enhanced by (a) intentionally locating learning communities within an institution's administrative structure, (b) dedicating line positions and hard budgets, (c) developing institutionally meaningful learning outcomes and a plan for reporting consistent efforts through ongoing assessment, and (d) ensuring that growth in learning community programs does not outpace growth in the infrastructure needed to support them. Finally, both faculty and staff involved in learning communities require—and deserve—continual training and development on a wide range of topics, including student and human development theory, educational psychology, and advanced pedagogies.

Future Considerations

The demand for postsecondary education to produce increasingly competent graduates is unlikely to abate; the opportunities and threats facing future generations seem to grow ever more complex with each passing day. To the extent to which these needs and the purported capacities of learning communities are coterminous and effective, learning communities will remain an enduring feature on the educational landscape for years to come. Increasingly, local program evaluations and multi-institutional studies suggest that their theoretical contributions can be realized. As has been noted, the development and implementation of learning communities are not simple, and they are not a panacea. However, as best practices from the field and the research continue to emerge, learning communities of all types can be strengthened to benefit

both students and faculty and, hopefully, their larger worlds.

*Matthew Soldner, Karen Kurotsuchi Inkelas,
and Aaron M. Brower*

See also Adult Learning; Cooperative Learning

Further Readings

Brower, A. M., & Dettinger, K. (1998, Nov/Dec). What *is* a learning community? Towards a comprehensive model. *About Campus*, 15–21.

Cross, K. P. (1998). Why learning communities? Why now? *About Campus*, *3*(3), 4–11.

Inkelas, K. K., Brower, A. M., Crawford, S., Hummel, M., Pope, D., & Zeller, W. J. (2004). *National Study of Living-Learning Programs: 2004 report of findings.* Retrieved October 18, 2006, from http://www .livelearnstudy.net/images/NSLLP_2004_Final_ Report.pdf

Lenning, O. T., & Ebbers, L. H. (1999). The powerful potential of learning communities: Improving education for the future. *ASHE-ERIC Higher Education Report*, *26*(6).

Levine Laufgraben, J., & Shapiro, N. S. (2004). *Sustaining and improving learning communities.* San Francisco: Jossey-Bass.

Shapiro, N. S., & Levine, J. H. (1999). *Creating learning communities: A practical guide to winning support, organizing for change, and implementing programs.* San Francisco: Jossey-Bass.

Smith, B. L., MacGregor, J., Matthews, R. S., & Gabelnick, F. (2004). *Learning communities: Reforming undergraduate education.* San Francisco: Jossey-Bass.

Tinto, V. (2000). Learning better together: The impact of learning communities on student success in higher education. *Journal of Institutional Research*, *9*(1), 48–53.

LEARNING DISABILITIES

According to the U.S. Department of Education, students with specific learning disabilities (SLD) account for nearly one half of students being served by special education under the Individuals with Disabilities Education Act (IDEA). There is a general consensus in educational and psychological literature that students with SLD have specific cognitive or psychological deficits that cause them to evidence academic deficits; however, the specific nature of these deficits has been highly debated. What is known, however, is that academic deficits that are a result of a learning disability

are hidden, often lifelong, and at times may be profound. Students with SLD generally demonstrate deficits in one or any combination of three basic academic skill sets: reading, written language, and math.

This entry begins with an introduction to the current federal definition of learning disabilities with a short discussion related to the evolution of the federal definition over the past 30 years. Next, characteristics of students with SLD will be presented; specific information on learner characteristics and challenges for each subgroup of SLD (reading, written language, and math) will be provided. Lastly, common service delivery models will be discussed, and information on educational implications for students with SLD will be provided.

Past and Current Definitions

Federal Definition

Specific learning disability was first defined under IDEA in 1975. The initial definition of SLD was as follows:

> a severe discrepancy between achievement and intellectual ability in one or more of the areas: (1) oral expression; (2) listening comprehension; (3) written expression; (4) basic reading skill; (5) reading comprehension; (6) mathematics calculation; or (7) mathematic reasoning. The child may not be identified as having a specific learning disability if the discrepancy between ability and achievement is primarily the result of: (1) a visual, hearing, or motor handicap; (2) mental retardation; (3) emotional disturbance, or (4) environmental, cultural, or economic disadvantage. (USOE, 1977, p. G1082)

Since then, many scholars have criticized the 1975 SLD federal definition for several reasons. First, the definition does not indicate that learning disabilities can, for some individuals, occur over the life span. Second, it does not indicate that students with SLD represent a heterogeneous group of students, each with unique academic deficits and learning needs. Third, and perhaps most important, the requirement of a discrepancy between IQ and achievement has been highly debated. Furthermore, research conducted over the past 20 years has shown that the discrepancy model has weak validity and reliability for identifying students with SLD.

The Achievement-Discrepancy Model of Identification

A main premise of the discrepancy model used in the determination of SLD is that the students who evidence a discrepancy between their IQ and achievement are distinct in cognitive processing than students who are solely low achievers in reading, math, or both. However, empirical research conducted since the 1975 federal definition has shown that there are no clear or significant differences in cognitive performance between discrepant students and those who are considered to be low achieving.

The use of cognitive assessments such as IQ tests have also been criticized by many scholars as inadequate, mainly because such approaches have not proven to be reliable or equitable across ethnic groups for correctly identifying students with learning disabilities. There has been little evidence to suggest that students who demonstrate a discrepancy between IQ and achievement have different developmental prognoses than children who are simply low achieving. Additionally, scholars have argued that focusing on IQ does not provide information about specific processing deficits or educational outcomes of students. Furthermore, IQ information does not provide relevant information regarding specific educational interventions for students who are having difficulties with reading or math.

Discrepancy models of identification have also been seen as inadequate in the way in which they have been implemented in the schools. Because of the nature of IQ tests, which in general require students to be of certain age, students are by and large not identified as SLD until middle or late elementary years. This model of identification has been called by many the "wait to fail" method, as it is difficult to identify students with a discrepancy between their intelligence and academic performance until they reach middle elementary grades. Once students reach these grades, it is difficult to implement programs that will help them "catch up" with their normal-achieving peers. Research has demonstrated that even very intense remedial interventions with this age group are not extremely successful. Not surprisingly, the "wait to fail" model does not readily allow for a closing in the achievement gap between normally performing students and those identified with learning disabilities. Instead, most students identified in middle elementary grades show very small gains in academic performance and remain in special education throughout their school career.

In sum, there are three critical problems associated with the IQ discrepancy model of identification:

1. Students are not identified early enough to receive effective intervention.

2. There are no meaningful differences between students who are identified with SLD and those identified as low achieving.

3. Intelligence tests do not provide information as to what intervention is necessary to remediate the students' reading difficulties.

Prevalence of Children With Learning Disabilities

Nationally, the occurrence of children identified as receiving special education services has been reported as slightly less than 5% of the total number of school-age children. It is important to note, however, that within the general population of students with disabilities, the largest category comprises students with a specific learning disability; this subgroup constitutes approximately 50% of all students with disabilities. Furthermore, the rate of growth for this category has grown tremendously with the number of students identified with SLD, nearly tripling since 1975.

Response to Intervention

In an attempt to address these critical issues within identification, response to intervention has been suggested as an alternative to the IQ discrepancy model for SLD identification. Frank Gresham has defined responsiveness to intervention (RTI) as the change in behavior or performance as a function of an intervention. Over the past 20 years there has been a growing body of research showing the usefulness of the RTI model as both a diagnostic tool for SLD and an early prevention model for students who may be at risk for reading failure. In general, RTI research has focused on reading disabled populations. Many researchers have identified RTI as a valid way to operationalize and diagnose learning disabilities, more specifically, reading disabilities.

In essence, the core of RTI is effective instruction in the general education settings and consistent progress monitoring to determine students' response to the provided instruction. "Treatment resisters" are

students who do not respond to the provided high-quality instruction and are ultimately students who, after a series of decision-making processes, are identified for evaluation for special education services. Students with reading disabilities have been targeted for this model because decades of reading research has shown that prereading skills, such as phonological awareness, alphabet knowledge, and decoding, are essential for successful reading achievement in later elementary grades. This extensive line of research has identified early skills necessary for reading, which makes reading development and reading disabilities identification a natural match for RTI, a method that is grounded in early intervention and prevention of future learning difficulties.

No Child Left Behind

The reauthorization of the Elementary and Secondary Schools Act (No Child Left Behind Act of 2001, P.L. 107–110) has created an instructional environment with successful reading as the primary goal of early elementary school. For example, the Reading First program requires implementation of scientifically based reading instruction within a multitiered model of instruction in general education settings. These policy changes, coupled with the reauthorization of the Individuals with Disabilities Education Act (IDEA), have led to a shift in the definition of learning disabilities, as schools are no longer required to use the discrepancy model of identification. Instead, the wording reflects the following:

> In determining whether a child has a specific learning disability, the LEA *shall not* be required to take into consideration whether a child has a severe discrepancy between achievement and intellectual ability. (Section 614 (b)(6A)).

Further, the act states:

> In determining whether a child has a SLD, the LEA *may* use a process that determines if the child responds to scientific, research-based intervention as part of the evaluation procedures described in paragraphs 2 and 3 (of Section 614). (Section 614 (b)(6B)).

The reauthorized IDEA allows RTI data to be used in the identification and eligibility determination of students with SLD, essentially as an alternative to the discrepancy model. In addition, it allows school districts to allocate more funds for early intervening services for students who are not eligible for special education, but who have educational or behavioral needs that qualify them as at risk for school failure.

RTI models align with Reading First, in that they are designed around a three-tiered model of instruction, with tier 1 being high-quality, general education instruction for all students; tier 2 including supplemental instruction for student who are struggling; and tier 3, an even more intense intervention for students who are still nonresponsive after tier 2 instruction has been implemented. Students may be considered for special education if they are dually discrepant, that is, if they are performing below the level of their peers and their learning rate is lower than that of peers.

Characteristics of and Challenges Faced by Students With Specific Learning Disabilities

Reading Disabilities

In general, reading disabilities are divided into two levels: word-level disorders and text-level disorders. Word-level disorders are further divided into decoding disorders and fluency disorders. Text-level reading disabilities are related to reading comprehension difficulties. Furthermore, it has been demonstrated that reading disabilities are persistent and different from developmental lag of reading skills seen in some children. Additionally, the majority of children (74%) identified as reading disabled in third grade remain disabled in ninth grade.

Decoding Disorders

Word-level decoding disorders are often called *dyslexia* and generally manifest themselves in the deficiency in a person's ability to decode single words. Empirical literature suggests that word-level decoding disorders are caused by difficulties in *phonological processing*, that is, the ability to understand speech sounds (*phonemes*) and process the individual sounds to make words.

In general, decoding deficiencies lead to profound difficulties with reading and reading comprehension. Furthermore, it is suggested that students who are not able to decode words with ease have a difficult

time comprehending text, as the majority of their concentration is dedicated to decoding words and not understanding the meaning of the text.

Fluency Disorders

Fluency is the ability to read words accurately and quickly. The National Reading Panel defined *reading fluency* as "the ability to read a text quickly, accurately, and with the right expression." When readers are fluent, most words are recognized automatically; therefore, fluency disorders are often seen as a consequence of decoding difficulties. However, there is research to suggest that fluency disorders can be present when decoding disorders are present. Explanation for fluency disorders in the absence of decoding disorders are usually related to poor attention or poor allocation of cognitive resources.

Fluency is often seen as the bridge between decoding and successful comprehension of text, because if students can recognize words rapidly and with ease, more concentration can be spent on inferring meaning from text.

Comprehension Disorders

Comprehension is a specialized type of reasoning that is conditioned and bound by the content and cognitive requirements of written text. Comprehension requires that students construct coherent representations of text by rapidly recognizing words, accessing a network of semantic relations associated with these words, and, guided by syntax, detecting or constructing meaningful relationships among words. Comprehension requires a complex set of processes beyond processes that are required for accurate decoding and fluent word reading. Although accurate comprehension assumes fluent word reading, it is not necessarily sufficient for successful reading comprehension. There is some research to suggest that some students evidence comprehension disorders with an absence of decoding and fluency disorders.

Written Language

Writing

The literature is clear that the writing process is composed of three main subprocesses: planning, sentence generation, and revision. Students with weaknesses in writing may exhibit additional difficulties in tasks such as outlining in a logical sequence, composing simple to complex sentences, and integrating word-level writing to sentence or paragraph forms. Additionally, children with a writing disorder may evidence extremely poor handwriting; this may be attributed to the inability to perform the requisite motor movements, also known as *dysgraphia.*

Spelling

Students with an SLD may experience challenges in their ability to correctly spell words. This difficulty may be due to students being unable to merge the relationship between spoken and written language to improve their literacy development. The skill set of poor spellers may be improved, however, with explicit instruction of letter patterns and multiple opportunities for practice.

Math Disabilities

A math disability, or *dyscalculia*, refers to students who evidence conspicuously poor skills at using basic computational processes to solve equations, including mathematical thinking and problem-solving skills. The ability to successfully execute each of these skills requires the ability to successfully manipulate multiple processes, such as retrieval in computation and language-based tasks involved in problem solving.

Metacognition

Metacognition refers to one's awareness of how one thinks and how one monitors what one thinks. Unlike typical learners, students with SLD often display difficulties in monitoring their learning and utilizing strategies or problem-solving skills. This inability to know a large number of strategies and understand how to use them efficiently is a characteristic often experienced by students with SLD.

Instructional Approaches for Students and Educators

The settings within which students with SLD are educated, and the implementation of effective instructional approaches used, have been an area of great controversy over the years. Settings within which students with SLD may be educated include inclusive classes, resource support programs, separate class placement, and separate school placement. Additionally, there are several instructional recommendations

that have been validated in the literature to specifically address and improve skill deficits for students with SLD. These settings and instructional recommendations will be discussed in detail according to group type (early intervention, elementary and middle school levels, and secondary and transition level), and suggestions will be provided to provide necessary information for the improved academic successes of those students who, with the right supports in place, can learn.

Service Delivery Settings

Service delivery settings lie on a continuum, which ranges from highly supported and separate (separate class or school) to highly integrated with the general education (inclusive class). Support for students with disabilities being educated for as much of the school day as possible is protected by the IDEA least restrictive environment provision. This policy has been further improved to mandate that students with disabilities not only be included but also make progress in the general curriculum. Placement settings include the least restrictive placement, resource room, separate class, and, finally, separate schools. The least restrictive placement (most inclusive setting) is identified in the *Reports to Congress* as a "regular class" or where special education services are delivered to a student for less than 21% of his or her school day. Further down the continuum are more restrictive environments that include settings in which a student receives services outside of a general education classroom for more than 21% of the school day. These include a "resource room" placement where students receive services outside the general education classroom from 21% to 60% of the school day. Next is a "separate class" setting. This setting includes students who receive services outside of a general education classroom for more than 60% of their school day. Lastly, there is a separate school category. Although there is some evidence that may support students with disabilities being educated outside a general education classroom with instructional methods that may not be readily available in a general education classroom, evidence strongly supports the idea that students with learning disabilities be educated in a general education classroom for most of their school days.

Current figures report the educational placement of students with SLD include 49% of students in regular class 80% to 100% of their time, 37% of students in regular classes 40% to 79% of their time, 13% of students in regular class 0% to 39% of their time, and the remaining 1% in residential, separate facilities, or home or hospital programs.

Instructional Approaches

Much work has been done in the past few decades to contribute greatly to what is known about how to educate students with SLD. Joseph Torgesen specified that special education differs from general education for students with learning disabilities when it is more explicit, intensive, and supportive. This includes instructional approaches that are characterized as being explicit, carefully designed, and closely related to the area of instructional need (e.g., reading, spelling, math). Furthermore, many instructional approaches have been found to improve the academic performance of students with SLD. These include embedded learning opportunities of key skills with multiple exposures and opportunities for engagement, differentiated instruction that involves actually differentiating educational content, and learning strategies instruction to provide students with SLD the means to become independent learners requiring fewer instructional supports. Other instructional approaches include scaffolded instruction, guided practice, modeling, reciprocal questioning, and feedback during instruction. Instruction can be further extended with the use of graphic devices, visual and verbal devices, and memory devices. Variations in instructional grouping that have been found to be beneficial include cooperative learning, peer learning, and structured small-group practice. Continual student monitoring is critical. The information collected during student monitoring of learning should be used to guide and deliver further instruction. Following is a brief description of these interventions or teaching methods that have been found to be beneficial for students with SLD.

Teaching Methods

Scaffolded Instruction

Scaffolded instruction is a process engaged in between teachers and their students whereby student understanding of a new concept is improved on through thoughtful dialogue directed by his or her teacher. Teachers who practice scaffolded instruction aim to extend their students' knowledge by beginning with highly teacher-mediated instruction of new

concepts or materials, always moving toward more independent or highly student-contracted learning.

Guided Practice

Similar to scaffolded instruction, guided practice focuses on the teacher acting as an educational coach and builds on a specific strategy or skill that has been described and modeled by the teacher. This is an opportunity for the teacher to shift the focus of instruction from being heavily teacher-mediated to more student-directed learning.

Modeling

Modeling is an opportunity for a teacher to use a "think aloud" process of instruction of a new concept or skill. This process is where the teacher, using student language, motors through the completion of the task being taught. This process allows the students to hear, in their own language, how a skill is initiated and completion is self-regulated. Essentially, it allows for students with SLD to see and hear how good thinkers think. Furthermore, it lends itself to providing greater opportunities for students with SLD to internalize the requisite processes to successfully complete a task.

Reciprocal Questioning

Reciprocal questioning is a process engaged in by two or more individuals where mutual learning is facilitated in a cooperative learning environment. This process allows for students to ask and respond to questions with each other to further their learning. It also allows for the teacher to provide individual feedback while the other students are engaged in their activities.

Feedback

Feedback is an approach that allows teachers to check for understanding of the group or with an individual. Typically, feedback is provided during an instructional phase during which students are actually working on completing a skill or task. With feedback, the teacher is providing necessary guidance through dialogue to facilitate learning by helping students get past an area of challenge. It helps students to become "unstuck" should they be completing a task incorrectly. The information gathered during feedback is critical to both student and teacher, as it also serves to help the teacher focus the instruction to ensure learning by all.

Devices to Improve Learning

Various types of devices can be used to enhance the teaching of a particular skill or concept. Devices have been used by effective teachers for many years and are helpful across different levels of students with SLD. Devices allow for a student to construct a deeper understanding of the information, but they are not enough when used in isolation. Teachers need to be explicit in their discussion of devices as simply a means to the end, the end being greater understanding of the underlying concept.

Graphic Devices

Graphic devices allow for students to see a visual representation of an abstract or complicated concept. Graphic devices are widely used. Common forms include maps, graphs, or Web diagrams to display cause and effect, comparison and contrast, problem and solution, hierarchical relationships, and sequence.

Verbal Devices

Verbal devices can also be used to enhance learning by making an abstract concept more concrete. This can be done by explaining the concept using an analogy to what is already familiar to the student. Using an analogy allows for the student to see the relationships between what is known and what he or she is attempting to learn.

Memory Devices

Memory-enhancing devices, or *mnemonics*, help students with SLD to remember and also retrieve necessary information. Common uses of these devices can include helping students remember lists of information and definitions of words. An example includes remembering the names of the Great Lakes with the mnemonic HOMES (Huron, Ontario, Michigan, Erie, and Superior). What is critical for success with memory devices is linking the mnemonic to information being learned.

Grouping Variations

Variations in traditional models of instruction have been shown to be effective for students with SLD. These models include cooperative learning, peer-assisted learning, and structured small-group instruction. Cooperative learning provides opportunities for

peer interaction and group participation by having students come together to practice skills learned. Classwide peer tutoring is a structured peer learning process. Each student has a role and is taught the responsibilities of the role of tutor or tutee. Specific content material is chosen by the teacher, and students are paired and work together in each of the roles. Small-group instruction has been shown to be beneficial for students in the early intervention model as well as those in elementary, middle, and high school. Evidence in these areas suggest that lower student-to-teacher ratios may improve learning outcomes for students with SLD. Furthermore, smaller groups may allow for better customization of instruction to the individual needs of the students involved because of reduced variability among the students.

Learning Strategy Instruction

Learning strategies are an effective way to teach students with SLD to become more effective learners. Essentially, strategy instruction is teaching students how to learn and then how to demonstrate what they know. Based on what is known today, educators may no longer assume that their students have the requisite sets of skills required to complete academic tasks before them. These tasks include taking tests; writing sentences, paragraphs, or essays; and organizing their information. There are three underlying principles of effective strategy instruction. These include, first, *cuing* a student to do something. Second, a strategy should provide a means to remember the steps required. Third, the strategy should address an area of difficulty for the student. For example, the paraphrasing strategy is a strategy designed to teach students to paraphrase written material for better comprehension. The strategy uses the mnemonic RAP (Read a paragraph, Ask yourself what is the main idea and what are the details, and Put the main idea and details in your own words). The sequence of the steps outlined in combination with the teaching procedures or stages of acquisition can help students become more independent in their reading and generalize the strategy use to other settings. Learning strategies help students with SLD become effective, efficient, and independent learners.

In sum, much work has been done to contribute to what we know about teaching students with SLD. It is important to note, however, that even with this knowledge, many students still do not respond to this instruction. This may be attributed to the rate and intensity of implementation or fidelity of the instruction. Providing supports for students with SLD can improve their ability to be successful at various skills and academic tasks and facilitate their independent learning. Furthermore, these academic supports can be implemented in a range of settings, the least restrictive being a general education-inclusive classroom.

Nanette S. Fritschmann and Emily J. Solari

See also *Diagnostic and Statistical Manual of Mental Disorders*; Disabilities; Individuals with Disabilities Education Act; Least Restrictive Placement

Further Readings

Bradley, R., Danielson, L., & Hallahan, D. (2002). *Identification of learning disabilities: Research to practice.* Mahwah, NJ: Lawrence Erlbaum.

Fletcher, J. M., Lyon, G. R., Fuchs, L. S., & Barnes, M. A. (2007). *Learning disabilities: From identification to intervention.* New York: Guilford Press.

Gresham, F. M. (2002). *Responsiveness to intervention: An alternative approach to the identification of learning disabilities.* Mahwah, NJ: Lawrence Erlbaum.

Lenz, B. K., Deshler, D. D., & Kissam, B. R. (Eds.). (2004). *Teaching content to all: Evidence-based inclusive practices in middle and secondary schools.* Boston: Pearson Education.

MacMillan, D., & Siperstein, G. (2002). Learning disabilities as operationally defined by schools. In R. Bradley, L. Danielson, & D. Hallahan (Eds.), *Identification of learning disabilities: Research to practice.* Mahwah, NJ: Lawrence Erlbaum.

McLeskey, J., Hoppey, D., Williamson, P., & Rentz, T. (2004). Is inclusion an illusion? An examination of national and state trends toward the education of students with learning disabilities in general education classrooms. *Learning Disabilities Research & Practice*, 19(2), 109–115.

No Child Left Behind Act of 2001, P.L. No. 107–110, 115 Stat. 1425 (2002).

Vaughn, S., & Fuchs, L. S. (2003). Redefining learning disabilities as inadequate response to instruction: The promise and potential problems. *Learning Disabilities Research and Practice*, 18(2), 137–146.

Vaughn, S., & Linan-Thompson, S. (2003). What is special about special education for students with learning disabilities? *The Journal of Special Education*, 37(3), 140–147.

Zigmond, N., Jenkins, J., Fuchs, L. S., Deno, S., Fuchs, D., Baker, J., et al. (1995). Special education in restructured schools: Findings from three multi-year students. *Phi Delta Kappan*, 76, 531–540.

Zipperich, M. (1995, January). Teaching web making as a guided planning to improve student narrative writing. *Remedial and Special Education*, *16*(1), 3–15.

LEARNING OBJECTIVES

The *learning objective*, clearly one of the most influential and useful concepts in educational psychology, may be defined as a statement placed within an instructional lesson that describes what the student should learn. The statement must be unambiguous and describe the desired instructional outcomes in terms that are specific and observable. The purpose of the learning objective is to help the student focus on those aspects of the material to be learned that will help the student successfully complete the lesson. A learning objective may also be called an *instructional objective*, *behavioral objective*, *performance objective*, or simply an *objective*.

Characteristics

The following is an example of a typical learning objective in the subject matter area of mathematics:

> The student will complete multiplication problems where each one is two randomly selected 2-digit numbers; the student will achieve a score of 80% or greater on a 20-item test. The student will work only with paper and pencil and must complete the test within 30 minutes; he or she is not permitted to use calculators, refer to multiplication tables, or receive assistance from another person.

Learning objectives share the following characteristics:

1. Learning objectives are designed for use in one lesson and are not to be viewed as broad goals, according to Robert Mager, one of the first advocates of learning objectives in education.

2. Learning objectives are not written for the teacher but for the student. Learning objectives do not describe teaching methods or media used in the lesson.

3. Learning objectives do not present information about the content of the lesson.

4. A learning objective has a number of critical components, and only an individual who has a good understanding of a lesson can write an effective learning objective.

ABCD Mnemonic

A convenient way to remember each part of a learning objective is to use the ABCD mnemonic.

A is for audience. When one reflects on this aspect, one should address the following questions: Who will use the objective? Do the students have the prerequisite knowledge needed to complete the lesson? Are the students ready developmentally for the material? In other words, the instructional designer must know how the lesson fits within the curriculum and the learner's capabilities.

B is for behavior. When considering this aspect, one addresses the following questions: What will the student be able to do once he or she has successfully completed the lesson? What should one expect to observe? The heretofore-mentioned learning objective for mathematical content avoids language such as "understands multiplication" or "knows multiplication," because such phrases are ambiguous. "Know" and "understand" describe mental states that cannot be observed, and consequently, it will be difficult for the student to determine what constitutes mastery performance. An advocate, such as Mager, of learning objectives stated in behavioral terms, would recommend that terms in a learning objective that describe mental states be replaced with language that describes observable behaviors. The prior learning objective in arithmetic should provide a precise definition of the type of problem that the student is expected to solve.

C is for conditions. This aspect of the learning objective relates to the context of the student's evaluation. In the earlier example of a learning objective in arithmetic, the student is required to solve the problems in 30 minutes without assistance. The "conditions" portion of the learning objective allows the student to gauge how difficult the evaluation will be and should influence how the student approaches the lesson.

D is for degree. That is, one indicates the level of performance that the student must attain to achieve mastery. In the arithmetic example, the student must reach 80% or greater on a 20-item examination. Without "degree" information, the student may mistakenly develop a skill level that is inadequate for the lesson goals.

Taxonomies

Not all advocates of learning objectives restrict learning objectives to behaviors devoid of mental terms. Some advocates such as Benjamin Bloom contended that behaviors described in a learning objective can include any type of skill or knowledge that is to be learned from a lesson. This viewpoint has informed the development of three taxonomies of educational objectives: one for cognitive educational objectives, a second for affective educational objectives, and a third for psychomotor educational objectives.

The most famous of these three taxonomies is the taxonomy of cognitive educational objectives written by Bloom and his colleagues. Cognitive objectives include different types of intellectual skills, such as memorizing, decision making, and reasoning. Bloom recognized that different types of cognitive objectives form a hierarchy. Arranged from simple to complex, this hierarchy consists of six classes of cognitive objectives: Knowledge, Comprehension, Application, Analysis, Synthesis, and Evaluation.

The taxonomy of affective objectives covers topics such as attitude change and emotional development. Affective learning objectives may be written for lessons on art appreciation, race relations, or ethical behavior. Categories of such objectives include attending, responding, and valuing.

The psychomotor category covers lessons on the development of bodily movements. This would include physical activities, such as sports, keyboard use, dancing, and tool use.

Additional Benefits

Although learning objectives are designed for the student, a well-written objective can also help the instructional designer or teacher improve instruction. For example, the objective can help one make more effective decisions about medial selection, instructional techniques, classroom evaluation, and remedial education. In other words, a learning objective reduces ambiguity for the teacher as well as the student.

Despite the usefulness of learning objectives, they should not be viewed as a requirement for all lessons. For example, J. M. Lewis cites research demonstrating that learning objectives allow the student to progress faster through lessons, but they do not consistently improve the student's academic performance. Lewis suggests that the instructional designer should consider other techniques to direct student attention in a lesson. Alternatives may include directions, pretests, advance organizers, and questions embedded in the lesson. But whether one views learning objectives as a requirement of all course lessons or not, learning objectives will continue to be useful and influential in the corpus of educational psychology.

Michael Petkovich and William M. Bart

See also Bloom's Taxonomy of Educational Objectives

Further Readings

Bloom, B. (Ed.). (1956). *Taxonomy of educational objectives; the classification of educational goals: Handbook 1. Cognitive domain.* New York: McKay.

How to write clear objectives. (n.d.). Retrieved November 29, 2006, from http://tlt.psu.edu/suggestions/research/Write_Objectives.shtml

Krathwohl, D. R., Bloom, B. S., & Masia, B. M. (1964). *Taxonomy of educational objectives; the classification of educational goals: Handbook 2. Affective domain.* New York: McKay.

Lewis, J. M. (1981, March). *Answers to twenty questions on behavioral objectives. Educational Technology*, 27–31.

Mager, R. F. (1962). *Preparing instructional objectives.* Belmont, CA: Fearon.

TenBrink, T. D. (2003). Instructional objectives. In J. W. Guthrie (Ed.), *Encyclopedia of education* (2nd ed., Vol. 4, pp. 1176–1178). New York: Thomson-Gale.

LEARNING STRATEGIES

For the purpose of this entry, *learning* is defined as behavior change that is governed by the feedback that the child receives from his or her environment. Hence, learning is a process that can be accelerated or decelerated depending on environmental conditions. The focus of this entry is to describe a paradigm within which learning can be evaluated and environmental conditions can be altered and optimized to accelerate learning for all children. Learning results from an interaction between the learner and the learner's environment can include a variety of influential variables, some of which result automatically when the learner responds (e.g., computer plays a song

when the correct answer is selected), some that are under the instructor's control (e.g., teacher-directed variables), and some that are under the learner's control (e.g., child adds a set incorrectly and when checking the sum against a set of manipulative materials learns that he or she counted incorrectly). The environmental changes that precede and follow learner behavior (both programmed or intended events and unintended events) serve to either strengthen or weaken the learned association. This entry describes a paradigm for integrating assessment of learner performance with specific strategies empirically demonstrated to accelerate learning.

Instructional Hierarchy

The link between assessment and learning has been well documented and endorsed as a primary characteristic of effective teaching models. Child performance may be characterized in terms that link directly to instructional techniques of demonstrated effectiveness for that particular performance. Therefore, performance-based assessment before instruction is essential to determine prior knowledge that each child brings to the learning task. Regardless of the philosophies that teachers and researchers bring to understanding and promoting learning, the integration of performance assessment and instruction has broad utility and applicability to learning and may be incorporated within all approaches to instruction.

It is helpful to view learning as occurring in stages: initial learning for accuracy or quality; practice for fluency and endurance; and application, or combination of the components into new responses, to solve novel or more complicated problems. Hence, the instructional hierarchy is a useful heuristic for organizing learner proficiency and therefore identifying which strategies are likely to accelerate (and decelerate) learning for that particular learner. The instructional hierarchy has been widely and successfully used to design, implement, and evaluate interventions for children's academic performance problems.

The hierarchy contains four stages of learning: acquisition, fluency, generalization, and adaptation. Knowing the stage at which a learner's skill falls relative to goals for learning communicates to instructional facilitators exactly which training goals, performance measures, and associated set of instructional strategies will be most effective for that learner. The instructional hierarchy

provides a framework for instructional decision making that is driven by learner competence. Specifically, the instructional hierarchy provides a system for translating a continuous variable like learning into categories of proficiency for a skill or series of skills drawn from the instructional objectives. The learner's performance or proficiency determines which instructional strategies will be most effective for that particular learner at that particular point in time. Hence, the instructional hierarchy provides a basis for implementing effective instructional strategies and for revising these strategies over time as learner skill improves (and new strategies are needed). Characteristics of learner performance and associated instructional strategies are described for each stage of the instructional hierarchy in the following sections.

Acquisition

The first stage of the instructional hierarchy is the acquisition stage. A learner's performance at the acquisition stage is characterized by low accuracy and subsequent dysfluency. When children are first learning a skill (acquisition stage of learning), responses will be hesitant and sometimes incorrect. The goal of acquisition-level training is for learners to perform a new skill or behavior accurately without assistance. Instruction is provided on a continuum ranging from less intensive and more naturalistic strategies to more intensive and less naturalistic strategies. Because performance is continuously monitored as the learner acquires a new skill, instructional facilitators or trainers might initially provide less intensive and more naturalistic instruction embedded into the typical environment. If progress monitoring does not reflect discernible acceleration of learning, more intensive instructional programming can be undertaken at this stage, which might include more directive, less naturalistic approaches such as direct instruction. Direct instruction has been demonstrated to be an effective core and supplemental instructional approach that reliably accelerates learning.

Certain instructional strategies are of primary importance when learner performance is at the acquisition stage of learning. For example, instructional feedback should immediately occur. Hence, during this stage of learning, independent practice of the skill is counterproductive and actually gives the student a chance to practice the skill incorrectly. Guided practice is the ideal form of instruction whereby a peer or

teacher monitors each learner response and provides immediate corrective feedback. The content of the feedback at this stage should be correct and quick (interrupting the incorrect response). Delayed feedback or error correction will be inefficient and ineffective at the acquisition stage of learning. In addition to immediate corrective feedback, certain antecedent strategies are important at the acquisition stage. For example, extensive use of cuing, prompting, and modeling of correct responses when the teacher anticipates an incorrect response is pivotal to establish correct responding. Effective instructional facilitators have a system of antecedent support that includes providing task materials of appropriate difficulty, providing structured exemplars of correct and incorrect responding, and providing increasingly intensive cues and prompts until the correct response occurs. Important to the successful use of these strategies is the adequacy of the learning environment, most often classrooms. To successfully program acquisition-level instruction, the teacher should maximize time devoted to active student engagement in the lesson (minimizing behaviors that interfere with academic engagement) and should briskly pace instruction while frequently checking for student understanding. Effective teachers have a system for tracking student responding when a new skill is introduced to ensure that learners have made the correct discrimination before moving to the next stage of learning where more independent learning can occur.

At the acquisition stage of learning, the goal is to establish correct responding with instructional support. The level of intensity of instructional support should be increased until correct responding occurs. Learners who fail to correctly respond given intensive prompting and feedback may require alteration of task materials (making the task easier) or additional instruction on prerequisite skills. Ongoing performance assessment is necessary to determine whether or not the learner has attained the goal of the acquisition stage of learning, that is, accurate and independent performance of the skill. Whatever antecedent and consequent supports have been provided to establish correct responding at this point should be removed to measure persistence of correct responding in the absence of instructional support. If accurate responding persists when instructional supports are removed, then the learner has reached the fluency-building stage of learning.

Fluency Building

Fluency is necessary to competency. *Fluency* is defined as the combination of accuracy plus speed and is typically converted to a rate-based metric, such as number of responses correct per minute. The temporal dimension of the response allows for a performance description that extends beyond the attainment of 100% accuracy. Fluency has been shown to be an element in the development of competency within a wide variety of situations including general and special education. Fluency also requires the ability to quickly link a component skill with other component skills to produce more complex learner responses. Fluent component or subskill performance enhances competency with complex skills. Fluent learners can identify which component skills will help them to solve more complex problem presentations without hesitation.

Learners whose performances fall in the fluency-building stage of learning are learners who can accurately respond when the response is required and can therefore tolerate longer intervals between instructional feedback occasions. Independent practice at this stage can be effective so long as materials are controlled such that the learner can continue to accurately and independently respond (i.e., low probability of errors). It is difficult to overstate the importance of adequately controlled task materials at each stage of the instructional hierarchy. At the acquisition stage, materials should be controlled such that cues for skill performance are salient to the learner. Hence, predictable and repetitive cues for responding are desirable. Variation is desirable only to the degree that the learner is given occasion to discern the conditions under which a response is correct and incorrect. In the fluency-building stage of learning, variation in task materials is desirable to the degree that learner errors remain low. Task materials systematically may be increased in difficulty at the fluency-building stage so long as learner performance remains accurate as speed of responding improves.

At this stage, feedback can be delayed and should be tied to motivating contingencies to maximize growth and learning. Reinforcing faster performance and blocking reinforcement for slower performance is the guiding principle. At this stage, feedback should not interrupt responding (because it interferes with fluency and is unnecessary). At this stage, feedback must still be correct, but it can be less elaborative.

Prompts and cues for correct responding should not be provided at this stage, because they interfere with fluency building. Accurate responding is established with frequent and immediate feedback, whereas fluency is built with frequent *delayed* feedback. The key is to provide feedback after short intervals of practice during which students are working to improve their performance. Feedback generally should be brief, rather than elaborative, preserving instructional time and maintaining a brisk pace of instruction. At this stage, overcorrection may be an effective consequence-based strategy that can be combined with delayed feedback. When corrective feedback is provided, the student is then asked to provide the correct response three times in quick succession. This strategy has been used successfully in reading instruction.

Similar to decision making at the acquisition stage of learning, ongoing assessment of learner performance is the basis for determining when the learner has completed the fluency-building stage of learning for a particular skill. Fluent responding governs a learner's ability to apply a learned skill to the solution of increasingly complex and novel tasks. Learners who can fluently perform a particular skill are ready for the maintenance and generalization stages of learning.

Maintenance and Generalization

Struggling learners generally fall into the acquisition or fluency-building stages of proficiency and, therefore, most often benefit from the strategies described for those two stages. However, learners operating in the generalization or adaptation stages may require interventions such as guided application of fluent skills under novel conditions and with more complex tasks. Another element necessary for a student to achieve competency is application. *Application* refers to learners being able to use a skill in a wide range of settings and situations, or to accurately discriminate between the target skill and similar skills that might be needed to contend with various task demands. The practice conditions that establish fluent responding are the same conditions that promote maintenance and generalization. Maintenance refers to the continued performance of the skill when instructional supports have been removed. Generalization refers to two specific events, the occurrence of the trained response under untrained conditions and the occurrence of different responses under the training conditions. Maintenance and generalization epitomize the goals of the learning effort. If learning strategies have been successful, then the learned skill will persist when the instruction ceases, will serve the learner in accessing new information or solving similar problems in the future, and will lead to generally greater adaptation for the learner in the future. After all, what good is instruction if it does not produce a robust and enduring skill set that has value to the learner over time? Fluent responding is a necessary pre-condition to maintenance and generalization. Maintenance and generalization is also assessed via ongoing performance assessment. Brief timed episodes of responding yielding a score of responses correct per minute may be utilized to assess the learner's skill trajectory during the fluency-building stage of learning. Similarly, responses correct per minute in the absence of instruction and over time provides important information about the persistence of a learned skill and whether re-training might be needed given certain goals of instruction. Research has demonstrated that greater fluency is associated with successful maintenance, faster recovery or return to fluency, and faster acquisition of related skills. Hence, training to fluency is an important goal related to establishing a skill that maintains and has generality for the learner. Maintenance of foundational skills should be verified to ensure successful learning of related more complex tasks in the future. Further, responding under conditions that differ from those provided during instruction and more typical of the everyday conditions under which the behavior will need to occur to effectively serve the learner should be examined. For example, students who learn to read a number of content-controlled passages at high fluency levels that maintain in the absence of reading instruction may not necessarily demonstrate adequately changed reading performance on a task in a science lesson that requires successful passage reading. The degree to which the learned skill (fluent reading) generalizes to conditions that are different from those encountered during training (e.g., reading a science lesson) or to responses that allow for successful performance under typically encountered conditions (e.g., looking up an unknown vocabulary word to comprehend a science lesson) can and should be directly assessed.

When a learner has demonstrated sufficiently fluent performance and has been provided with reinforcement to apply that performance under different contexts and yet application does not occur, then the learner requires instructional support for generalization. Instructional facilitators should provide guided

practice opportunities to cue the need for application, assist the learner to apply the skill in that context, and provide immediate feedback to promote correct application. If needed, the instructional facilitator may provide many opportunities to respond, applying the learned skill to novel or more complex problems to build fluency for application of the learned skill. So, for example, often learners will successfully apply their knowledge of basic sums and subtraction facts 0 to 20 to the solution of word problems that utilize those computations, fact family problems, or multidigit addition or subtraction without, or even with, regrouping. For learners who do not automatically apply the learned skill, the same progression of instructional support to build accurate and fluent application may be needed to support successful application of the learned skill.

Adaptation

Even when a student is able to use a skill in many situations, he or she may not yet be able to modify or adapt the skill to fit novel task demands or situations. The goal of attaining competency is for the student to become capable of identifying elements of previously learned skills that he or she can adapt to new demands and situations. So, for example, a child who is at the adaptation stage may readily modify the skill of basic fact of $10 + 20$ to solve $10 + 21$, by solving the basic fact and adding one to the solution. Adaptation can be supported by prompting, suggesting, or reviewing possible strategies for students at this stage of instruction; however, this should only occur for learners who have surpassed the generalization stage of learning. Instruction often inadvertently occurs at this level for learners who have not demonstrated accurate, fluent, and generalized performance, and this mistake of instruction leads to errors in performance and decelerated learning trajectories in general.

Systems of Data-Based Decision Making to Facilitate Learning

The use of routine performance assessment is fundamental to identifying instructional strategies that will accelerate learning. Instruction that is not matched to student proficiency will be less effective and will negatively affect learning over time. Those who wish to facilitate learning must have a system for monitoring student performance and responding to those data in an iterative fashion to support continued learning toward competency. In addition to using routine performance assessment linked to specific instructional strategies, researchers must attend to the technical adequacy of the dependent variable (measurement) and the independent variables (implementation and measurement). Recently, systems of data-based decision making have been designed to translate learner performance information into instructional targets and measure the effect of those instructional efforts to reach decisions to ultimately improve learning for all children. These systems of data-based decision making have been referred to as *response to intervention (RTI)* models. Use of RTI models in schools has resulted in more equitable identification of children requiring special education services, lower numbers of children requiring special education, and improved learning for all children, most particularly for those at risk for failure. RTI evolved out of a desire to utilize the information obtained from routine performance assessment and instructional response to reach judgments of importance in schools, including who should receive supplemental or special education and how existing resources could be used to most efficiently accelerate learning for the greatest number of students possible. RTI is a very promising science of decision making that is being widely used following recent revisions to educational policy in the United States. Yet, the effectiveness of RTI (or the degree to which routine learning assessment is linked to use of strategies that enhance learning) depends upon two critical variables: intervention integrity and decision-making accuracy.

Intervention Integrity

Understanding how (and how well) an intended intervention was carried out is critical to internal, external, and consequential validity. Arranging antecedent conditions is a necessary but not sufficient condition to promote integrity of intervention implementation. Research is clear in suggesting that implementation integrity cannot be assumed. Rather, integrity of instructional implementation also must be routinely assessed and high levels of integrity must be strategically programmed. Research has found that to ensure adequate implementation of desired instructional strategies, teachers should be provided with

needed materials and should receive an in vivo performance coaching system that uses a combination of verbal instruction, written prompts detailing intervention steps, modeling, coaching, and ongoing performance feedback about the accuracy of instructional implementation along with documented changes in learning associated with correct use of the instructional strategy.

Decision-Making Accuracy

Decision rules should be operationalized, linked to data-collection procedures, and monitored for accuracy. Researchers have found that decision-making teams in schools rarely reach decisions about the need for a special education eligibility evaluation that match with (a) their own assessment data, (b) others' assessment data, (c) or even local criteria for identification. Hence, monitoring and quantifying the degree to which decision makers correctly utilize performance data to reach decisions about what types of instructional strategies are needed and when children might require more intensive support, such as that provided through special education, is essential.

Amanda M. VanDerHeyden

See also Applied Behavior Analysis; Curriculum Development; Direct Instruction; Individuals with Disabilities Education Act; Learning Disabilities; Measurement; No Child Left Behind; Precision Teaching

Further Readings

Burns, M. K., VanDerHeyden, A. M., & Jiban, C. (2006). Assessing the instructional level for mathematics: A comparison of methods. *School Psychology Review*, *35*, 401–418.

Daly, E. J., III, Lentz, F. E., & Boyer, J. (1996). The instructional hierarchy: A conceptual model for understanding the effective components of reading interventions. *School Psychology Quarterly*, *11*, 369–386.

Fuchs, L. S., & Fuchs, D. (1986). Effects of systematic formative evaluation: A meta-analysis. *Exceptional Children*, *53*, 199–208.

Haring, N. G., Lovitt, T. C., Eaton, M. D., & Hansen, C. L. (1978). *The fourth R: Research in the classroom.* Columbus, OH: Merrill.

Johnson, K. R., & Layng, T. V. J. (1992). Breaking the structuralist barrier: Literacy and numeracy with fluency. *American Psychologist*, *47*, 1475–1490.

Kameenui, E. J., & Carnine, D. W. (1998). *Effective teaching strategies that accommodate diverse learners.* Upper Saddle River, NJ: Merrill.

Lentz, F. E., & Shapiro, E. S. (1986). Functional assessment of the academic environment. *School Psychology Review*, *15*, 346–357.

Martens, B. K., & Witt, J. C. (2004). Competence, persistence, and success: The positive psychology of behavioral skill instruction. *Psychology in the Schools*, *41*, 19–30.

Vellutino, F. R., Scanlon, D. M., & Tanzman, V. S. (1998). The case for early intervention in diagnosing specific reading disability. *Journal of School Psychology*, *36*, 367–397.

Wolery, M., Bailey, D. B., & Sugai, G. M. (1988). *Effective teaching: Principles and procedures of applied behavior analysis with exceptional students.* Boston: Allyn & Bacon.

LEARNING STYLE

Learning style is a term used to portray individual differences in the way that people prefer to learn. Learning styles are typical patterns individuals use to process information or approach learning situations. These learning style preferences are thought to occur naturally. According to learning style theory, when an individual's learning preferences are met, the individual learns more easily and effectively. There are more than 70 theories and models of learning styles. Each model describes how particular kinds of individual differences influence learning. However, the kinds of individual differences and the ways that these differences influence learning vary considerably among theories. Although it is not possible or useful to summarize all theories, a few models that are frequently cited and described in the learning style literature are overviewed here and used as examples. Cautions and guidelines for using learning style information will also be discussed.

Importance

An underlying assumption of learning style theories is that individual differences in learning preferences are positive and useful and that both learners and teachers will benefit from becoming aware of, and understanding, learning style information about themselves and others. Also implicit in these theories is

the idea that it is important to acknowledge learner diversity and to customize and individualize learning so the needs of all learners are met. When learners understand how they prefer to learn, they can seek preferred learning settings and learn to cope with settings that do not align to their learning preferences. Teachers, administrators, and program planners can use learning style information when planning teaching strategies and learning activities, evaluating learning, and developing programs and curricula. Teacher education about learning styles will help teachers recognize their preferred learning and teaching styles and provide alternative ideas and strategies to help teachers incorporate multiple teaching strategies to facilitate the needs of individuals with different learning styles.

Many models assert that the process of self-awareness of how one learns may be the most useful part of learning style information. In many models learning style information is seen as a key component in "learning to learn." When learners understand how they prefer to learn and find ways to incorporate their learning preferences into a variety of learning settings, they can become more effective learners. Thus, the purpose of learning style information is to increase learner awareness about how they prefer to learn, so that they can learn more effectively. Learners who understand their own learning styles can also recognize and seek learning situations that match their preferences and learn to develop skills and strategies for success in conditions that do not match their learning preferences.

Are Learning Styles Modifiable?

Individual differences in learning style can be seen as preferences or as stable traits. In most learning style models, a learning preference is not thought to be the same as a trait or ability. However, in some theories, learning styles are thought to be relatively fixed characteristics or traits inherent within individuals. When learning styles are seen as relatively stable or unchanging, the theory's emphasis is on the need for teachers and trainers to adapt learning strategies and settings to accommodate the needs of learners with different learning styles. Within these models, it is often proposed that it is essential for teachers to be aware of all learning styles and to be able to accommodate them. Teachers are encouraged to

present the same information in multiple formats and to provide optional learning assignments and options.

However, matching learning to the learner's style is not always possible. Most theories tend to see learning style as somewhat adaptable and assert learners can succeed in a variety of learning tasks, settings, and situations. In these models, learners are encouraged to identify their learning preferences and utilize strategies aligned to their learning style to maximize their learning. Learners are also encouraged to identify learning situations that are challenging and then develop learning skills and strategies to manage their learning in these situations. In models that conceptualize learning style as modifiable, often the focus is on becoming aware of preferences while, at the same time, recognizing and learning to adapt to learning tasks and situations that do not align with an individual's learning style.

Some models place learning styles within steps or stages of a learning process. In these models, learning is usually conceptualized as a multistep process. Learning style is seen as a preference for a particular step within the learning process. Because learning requires all steps in these models, learners are encouraged to expand and develop learning skills and strategies to complete the nonpreferred parts of the learning process.

Consensus indicates that learning style describes natural and innate preferences for learning. However, the literature also makes it clear that learning style preferences are not to be seen as limiting or defining. Everyone can learn to thrive in nonpreferred learning situations, as necessary. However, when natural learning preferences are identified and supported, learning is easiest and most comfortable.

Learning Style Models

There have been various attempts made to organize learning style models into groupings. Some classification systems organize models on the basis of how stable or flexible learning styles are thought to be. Other classifications are made based on the various influences on learning, such as learning modality preferences, personality preferences, social or emotional aspects, and preferences for particular cognitive processes. Still other models focus on the process of learning and individual preferences for engaging in parts of the learning process.

To simplify a complex topic, three general kinds of learning style theories will be overviewed: those focusing on input of information, which will cover preferences for sensory input and learning environments; those focused on the cognitive processing and organization of information; and those developed around a learning process. Although there are more than 70 different models, only a few of the most often-cited and well-established models will be discussed.

Perceptually and Environmentally Based Learning Style Theories

Some theories propose that sensory input or environmental preferences are key aspects of learning style. These theories focus on preferences such as the sensory modality in which information is provided, preferred activity level of the learner, and specific characteristics of the physical learning environment that may impede or enhance learning.

One popular way of looking at learning style is by preference for visual, auditory, tactile, or kinesthetic sensory modes. Seeing, hearing, touching, and doing may be more or less preferred by an individual learner. It is assumed that individuals will learn more easily and remember better if information is presented in one or two of these modes. Visual learners will prefer written or visual instructions and will observe and remember visual resources. Auditory learners will prefer verbal instructions, find discussion helpful, and remember what is said. Tactile learners will want to touch and manipulate learning materials and will remember best when they can handle materials. Kinesthetic learners will prefer hands-on active experiences and will remember best when they can move around or become physically involved with learning materials.

A learning style model by Rita Dunn and Kenneth Dunn builds on this basic sensory preference model and suggests there are at least 18 basic elements that influence how well an individual will take in and retain information. These elements are linked to four modalities: environmental, emotional, sociological, and physical. The physical modality includes the four sensory modality preferences (auditory, visual, tactile, and kinesthetic) as elements. Their model also includes three other physical modality elements: preferences for intake of food or drink, time of day, and mobility. Environmental elements include noise, temperature, and design of learning space. Sociological elements include preferences for learning with others, presence of authority, routines, and need for approval of others. Emotional elements include motivation, responsibility, persistence, and need for structure. In this model, all of these elements combine to form an individual's learning style. When these elements are in place, learning will be enhanced.

Some learning style models explore other kinds of perceptual differences among learners. Herman Witkin explored how an individual's perception was influenced by the context or field the information is contained in. He found two kinds of perception: field-dependent and field-independent. Field-dependent perception learners are influenced by the visual field in which visual information is presented, whereas field-independent learners are not. He then asserted field-dependent learners will perceive information globally and will have a social orientation to the learning. Field-dependent learners will thrive in more personal learning settings and when learning information that is personally relevant. Field-independent learners will perceive information analytically and will have an impersonal orientation to learning. They will thrive in learning settings that provide opportunities to ask questions, solve problems, and work independently.

Cognitive Processing Models

Many learning style models describe differences in the way a learner prefers to process information. Sometimes these models refer to cognitive processing styles rather than learning styles. The term *cognitive style* is used in various ways in the learning style literature. Cognitive styles are sometimes conceptualized as a part of an overall learning style; when thought of this way, it describes how individuals approach cognitive tasks rather than how they approach learning overall. However, in many models the terms *cognitive style* and *learning style* are used interchangeably. Similarly, in different models, the terms *thinking styles* and *intellectual styles* are also sometimes used to describe learning style preferences, as well as alternatively proposed as a component of learning styles.

In general, cognitive processing models of learning styles propose dichotomies of individual differences. These models usually acknowledge that both sides of a processing dichotomy are important parts of the learning process. They assert that individual learners will prefer one side of each dichotomy and will find it

easier to learn when this preference is accommodated in the learning situation. These differences are given a variety of names and are described in different ways, depending on the specific model; generally, however, they fit into three dichotomies: abstract versus concrete, sequential versus simultaneous, and action versus reflection. Models that incorporate social aspects of learning include an additional dichotomy: logical versus personal.

Learners with a concrete learning style prefer to learn about relevant, practical, and useful information. They are drawn to learning that can be immediately applied in their current situation. Learners with an abstract learning style prefer learning about ideas and concepts. They are drawn to theory and enjoy linking and connecting ideas. Sequential learners prefer a step-by-step approach to learning, whereas simultaneous learners prefer to process information in a nonordered, more random way. Simultaneous learners tend to focus on more than one fact or concept at a time, whereas sequential learners tend to take on one piece of information at a time. In some theories, simultaneous processing is also referred to as *random processing.*

Action-oriented learners prefer to interact with others and with materials when they are learning; they are stimulated by discussions and other opportunities to share ideas and information with others. Reflective learners, on the other hand, prefer to have time to think about material. They process information most effectively when they are alone. Learners who prefer a logical approach to learning are described as objective and tend to prefer frank, corrective feedback from a competent expert. Individuals with a values-based approach prefer cooperative and collaborative learning settings and like to receive positive feedback and support from an approachable, personable teacher.

Anthony Gregorc, in his mind styles model, proposes combinations of two processing preferences to delineate four learning styles. Gregorc uses the concrete versus abstract dichotomy to describe how learners perceive and understand information. He then adds a second dichotomy—sequential versus random—to describe how learners organize and order information. When these processing preferences are combined, his model results in four distinct learning styles: concrete sequential, abstract sequential, abstract random, and concrete random. An emotional and logical dichotomy is also incorporated into the model. For example, an abstract sequential learning style is characterized as logical and analytical, whereas an abstract random learning style is characterized as sensitive and emotional. An individual may have one strongly preferred learning style or may learn best with a combination of two or more styles.

Theories built on the personality type theory of Carl Jung link abstract and simultaneous processing. In personality type models, this preference is referred to as Intuition: a preference for taking in abstract information about ideas, theories, and concepts. This theory also links preferences for concrete and sequential learning. In the personality type model, this preference is referred to as Sensing: a preference for taking in information gained through the senses that is related to real-world experiences and applications. Personality type theory also proposes dichotomies between action and reflection (Extraversion and Introversion) as preferred orientation to the world, between logical and personal information processing (Thinking and Feeling), and between a structured and flexible approach to the environment (Judging and Perceiving). Because this model uses four dichotomies, the various combinations of preferences lead to 16 distinct learning styles with individualized learning preferences. Some learning style models simplify personality type theory and use differences in two of the four dichotomies to describe four learning styles.

Left- and right-brain models provide another way of looking at cognitive processing. In these models, the processing differences are verbal versus visual-spatial, sequential versus holistic, and logical versus emotional. Building on the idea that there are biological differences in the way the different sides of the brain process information, these models assert that the left side of the brain prefers to process information verbally, sequentially, and logically, whereas the right side of the brain processes information in a visual-spatial, holistic, and emotional manner.

Ned Herrmann created a learning styles model based on mental preferences or thinking styles from right- and left-brain models. In his model, there are four styles: the Theorists (rational), Organizers (safe keeping), Innovators (experimental), and Humanitarians (feeling). The Theorist learning style is characterized by preferences for logical analysis. Sequencing and structuring information characterizes the Organizer learning style. The Innovator learning style is characterized by preferences for exploration and self-discovery, and the Humanitarian learning style is linked to emotional involvement and harmony. In his

model, Herrmann asserts that it is difficult for learners to accommodate information presented in opposing styles. For example, a learner who prefers an Organizer learning style may have difficulty accommodating Innovator styles, or a learner with a Humanitarian learning style may find it difficult to accommodate a Theoretical style.

Learning Process Models

Some models, such as the one by David Kolb, propose that learning occurs in a process or cycle. Each part of the learning cycle requires the use of different combinations of processing preferences. In his model, the processing dichotomies (abstract vs. concrete and action vs. reflection) are not only linked to individuals' learning preferences but also incorporated into a defined learning cycle. In his model, the learning process begins with a concrete experience. Then reflective observation occurs, followed by abstract conceptualization and active experimentation. Learners tend to prefer one part of the learning process to the other. As a result of these preferences, four learning styles are proposed: abstract conceptualization combined with active experimentation (converging learning style), concrete experience combined with reflective observation (diverging learning style), abstract conceptualization combined with reflective observation (assimilating learning style), and concrete experience combined with active experimentation (accommodating learning style).

Peter Honey and Alan Mumford built on Kolb's work to identify four different learning styles. They proposed that learning comprises four steps: experiencing, reflecting, concluding, and planning next steps. Individuals tend to prefer one or two steps in the process over the others, so the four steps in the learning cycle link to four learning styles. The four styles include activists, who prefer to actively engage in experiences when learning; reflectors, who prefer to review or ponder experiences; theorists, who analyze and make conclusions; and pragmatists, who plan next steps. Similarly, Bernice McCarthy elaborated on Kolb's model. She linked the abstract versus concrete and action versus reflection processes to left- and right-brain functions. McCarthy proposed four learning styles within these new dimensions. In these process-based learning style models, learning styles are thought to be modifiable. Effective learners need to utilize all four

steps. Learning style information helps them understand what steps they prefer and provides insights into how to develop underutilized steps.

Measuring Learning Styles

From the various theoretical models, there are at least a hundred inventories developed to measure learning styles. Many of these instruments have not been subjected to stringent test construction, and because of this, some practitioners and researchers question the use of learning style inventories.

Many of these inventories do not have demonstrated repeatability or accuracy, so it is important for practitioners to assess the reliability and validity evidence provided with a learning style measure to determine if it is based on sound psychometric principles. Even when a reliable and valid instrument is used, most instruments have not yet demonstrated a clear link between learning style and an individual's ability to learn certain kinds of information or thrive in specific learning environments. For example, if a learner does succeed in an environment that is not aligned to his or her learning style, it is difficult to know if this success indicates that learning styles are not important. Such a finding may alternatively indicate the learner has been able to adapt to that setting. Researchers cannot easily know if learning was more difficult or uncomfortable in that situation than it would have been in another environment.

The plethora of terminology and the large number of models and inventories contribute to the difficulty of surveying and assessing learning style inventories. As with many kinds of applied research, multiple interest groups in various settings, including academics, schools, and businesses, are developing learning style measures. These various interest groups have different aims and approaches, and there is little integration of the information gathered about the instruments used in each sector. There is also a large interest in the commercial applications of learning style inventories. A commercial focus can lead to research and promotional material that is focused on marketing a product rather than demonstrating the reliability of the inventory or the validity of a model.

Many learning style inventories use a self-report format. Self-report instruments are subject to error if learners are unaware of, or unable to, accurately report their preferences. If learning style measures are

to be a tool for self-awareness, those taking them must be able to use metacognitive skills to accurately evaluate themselves. When using these kinds of measures, it is always important to use the results only as a guide and explore possible reasons for inaccurate results. Some administration instructions for using learning style inventories do include a self-validating process whereby learners think about, and can override, results on an inventory. However, it is possible that not all learners have the skill to engage in this kind of reflective, self-awareness process.

It is also important to ensure the results of learning style indicators do not limit or stereotype learners so as to discourage them or have them dismiss certain ways of learning. For example, some information is more abstract than concrete. If a learner believes he or she cannot learn abstract information because he or she has a concrete learning style, then the learning style information has limited rather than benefited the learner. It would be more useful to show the learner how to ground the abstract information in a concrete way or to teach other similar learning strategies for thriving in this situation.

It is possible to use observational data and metacognitive questioning as well as inventories to help students become aware of their learning preferences. For example, Barbara Given has created a methodology for assessing learning styles that incorporates information from many models within an observational and self-report format. This model considers emotional, social, cognitive, physical, and metacognitive aspects of learning preferences.

Validity of Learning Styles

With all of the different theories, models, and inventories of learning styles, it is difficult to know which ones are valid. Likely there are a number of individual differences and preferences in the way a learner absorbs and retains information and interacts with others in a specific learning situation that affects the ease of learning and comfort the learner experiences. Reviews of learning style models and instruments conclude that there is a lack of definitive evidence that would demonstrate the adequacy and validity of any one approach to learning styles. Reviewers often mention the importance of considering the context and purposes for learning style information when assessing and choosing to use a particular model or instrument.

When critiquing the learning style literature, it is important to understand the connotation of the labels and terminology used. For example, the word *abstract* is used in several learning style models. However, what is meant by *abstract* and the characteristics of learners who prefer "abstract" learning can vary greatly. To use learning style information, practitioners need to be able to sort through the multiple models, terminology, and concepts associated with learning styles to find a model that works in their situation. This can be accomplished by critically assessing the strengths and limitations of specific models. For example, if the purpose of learning style information is to help individuals understand how they prefer to learn, models with defined support resources and applications will be more useful than those that do not provide this link.

It also is important to ensure learning style models are not too limiting. This can be the case when learning styles are conceptualized as "either/or." For example, it may be unwise to characterize a learner as experiential if this infers they are not reflective. This kind of characterization might indicate an experiential learner cannot reflect or will struggle in situations that require reflection. Such categorization has not been demonstrated and may provide a mind-set in which learners unnecessarily avoid or feel inadequate in certain learning situations.

Some researchers propose it would be most effective to take an integrated approach to defining and describing learning styles by including a number of models and dimensions for learners to consider as they are thinking about and identifying their learning preferences. Such models would include sensory preferences, environmental preferences, instructional preferences, social and emotional learning aspects, and thinking or cognitive style preferences. Ultimately, learning style information is a tool to be utilized carefully by teachers, educational planners, and learners, with an emphasis on an increased awareness of learning and teaching preferences, on learning how to adapt to and manage all learning situations, and on enhancing learning and teaching effectiveness. Perhaps it is the process of helping students and educators to think about and determine their learning preferences that is more important than any specific model used to carry this activity out.

Donna Dunning

See also Cognitive and Cultural Styles; Learning Strategies; Metacognition and Learning; Teaching Strategies

Further Readings

Bostrom, L., & Lassen, L. M. (2006). Unraveling learning, learning styles, learning strategies and meta-cognition. *Education & Training, 48*(2–3), 178–189.

Cassidy, S. (2004). Learning styles: An overview of theories, models, and measures. *Educational Psychology, 24*(4), 419–444.

Coffield, F., Moseley, D., Hall, E., Ecclestone, K. (2004). *Learning styles and pedagogy in post-16 learning. A systematic and critical review.* London: Learning and Skills Research Centre.

Coffield, F., Moseley, D., Hall, E., Ecclestone, K. (2004). *Should we be using learning styles? What research has to say to practice.* London: Learning and Skills Research Centre.

Delahoussaye, M. (2002). The perfect learner: An expert debate on learning styles. *Training, 39*(5), 28–36.

Desmedt, E., & Valcke, M. (2004). Mapping the learning styles "jungle": An overview of the literature based on citation analysis. *Educational Psychology, 24*(4), 445–464.

Evans, C., & Sadler-Smith, E. (2006). Learning styles in education and training: Problems, politicisation and potential. *Education & Training, 48*(2), 77–83.

Guild, P. B., & Garger, S. (1998). *Marching to different drummers* (2nd ed.). Alexandria, VA: Association for Supervision and Curriculum Development.

Hall, E., & Moseley, D. (2005). Is there a role for learning styles in personalised education and training? *International Journal of Lifelong Education, 24*(3), 243–255.

Reid, G. (2005). *Learning styles and inclusion.* Thousand Oaks, CA: Sage.

Zhang, L., & Sternberg, R. J. (2005). A threefold model of intellectual styles. *Educational Psychology Review, 17*(1), 1–53.

LEAST RESTRICTIVE PLACEMENT

The term *least restrictive placement* refers to the concept of the least restrictive environment (LRE), a legal principle ensuring the social integration of people with disabilities. Social policies in the United States, including the Americans with Disabilities Act, require public entities to administer their programs in the most integrated setting appropriate to the needs of qualified individuals with disabilities to facilitate their positive relationships with nondisabled people in typical community settings. The focus of the LRE principle is not on the type of facility but rather on whether a placement can provide the necessary services an individual needs to have life experiences that most community members find desirable. Placements that might be appropriate for an individual receiving public health services range from institutional settings to noninstitutional residential settings, such as assisted living, home, and community-based settings. This entry is intended to clarify the LRE principle in making placement decisions in public health agencies, educational institutions, and human services.

From a legal perspective, institutionalizing a person with a disability who can benefit from living in the community, and who wishes to do so, constitutes discrimination, because removal to a separate setting severely limits the individual's ability to interact with family and friends, to work, and to live independently. The right of people with disabilities to receive services in the most integrated setting is not absolute, however. Least restrictive placement policies do not require measures that would fundamentally alter the nature of a public entity's programs, and costs may be considered in making that determination. In reviewing the fairness of placement decisions, courts consider not only the cost of providing care in integrated settings to individuals but also the range of services provided to others with disabilities and the state's obligation to offer those services in an equitable manner. Understanding the LRE principle, and applying it correctly, is especially important to family members, psychologists, and other professionals responsible for making special education placement decisions on behalf of students with disabilities.

Understanding the Principle of the Least Restrictive Placement

The principle of the least restrictive placement is one of six core principles of the Individuals with Disabilities Education Act (IDEA), the federal legislation guiding special education policy in the United States. The LRE principle stipulates that, to the maximum extent appropriate to their individual needs, students with disabilities are to be educated with nondisabled students. Special classes, separate schooling, or other removal from the general education environment is to occur only if the nature or severity of a student's disability is such that education in regular classes with the use of supplementary aids and services cannot be achieved satisfactorily. The purpose of the IDEA is to

ensure that students with disabilities receive a free, individually appropriate, public education (FAPE) emphasizing special education and related services that meet their unique educational needs and that prepare them for further education, productive employment, and independent living. The LRE principle supports the goals of the IDEA by ensuring that a special education student's Individualized Education Program (IEP) is implemented in a learning environment that meets the child's needs and that least restricts his or her social integration with typically developing peers.

In the language of public policy, the LRE principle is considered to be a rebuttable presumption, which means the law presumes that the regular class is the least restrictive placement for any student to receive appropriate instruction. The presumption of regular class placement is rebutted by convincing evidence that an alternative placement would provide a particular student with an appropriate education. For this reason, school systems are required to make a full continuum of alternative placements available for students that range from regular classes, special classes, separate schools, residential facilities, hospitals, and home settings. School systems are also required to make provision for supplementary services, such as resource rooms or itinerant instruction to be provided in conjunction with regular class placement.

In addition to requiring school districts to provide a continuum of alternative placements, the IDEA requires special education placement decisions to be made by a group of people, including the student's parents and others who are knowledgeable not only about the child but also about the meaning of the evaluation data and the placement options. The IDEA identifies five factors that decision makers must follow to ensure that individual students are placed in LREs based on their specific learning needs. Placement decisions (a) must be made annually, (b) must be based on the student's IEP, and (c) must give consideration to any potential harmful effect on the child or to the quality of the required services. Unless their individualized programs require otherwise, (d) students with disabilities are expected to attend either a school as close to home as possible or the school they would attend if they were not disabled, and (e) they must not be removed from education in age-appropriate regular classrooms solely because of needed modifications in the general curriculum.

The least restrictive placement, or environment, is not a specific location but the outcome of a procedural process in which a greater weight is given to the standard of appropriate services based on individual needs than to the factors that determine the placement in which the services occur. In making placement decisions, courts compare the student's educational needs with the statute's overall preference for placement in the regular classroom. Parents and professionals, less familiar with the law's presumptive language, often confuse the terms *mainstreaming, inclusion*, and *LRE*, which do not have interchangeable meanings. Mainstreaming implies that special education and general education students will be educated together as appropriate but not exclusively. Full inclusion implies that students with disabilities have an absolute right to regular class placement. The term *LRE* is not synonymous with *inclusion* but rather requires that placement decisions for students be made on the basis of an appropriate IEP. LRE is a legal principle embodied in the IDEA; in contrast, the practices of mainstreaming and inclusion are not required by law but are educational strategies used to make the least restrictive placement principle operational in schools.

Where a student receives instruction is only one component of an appropriate education. Some students may need instruction that cannot be provided in regular classes because they need to learn something different than general education students, such as Braille, American Sign Language, or specific technologies that are more efficiently taught in other settings. Some students may need to learn things differently, such as students with severe learning disabilities who need intensive reading instruction in more private learning environments, or students with cognitive disabilities whose job training and life skills curriculum requires them to spend time in community-based settings. The LRE requirements compel school systems to make a continuum of options possible so that IEP teams can make appropriate student-centered placement decisions across a range of viable alternatives. With regard to special education placements, decision makers need to not only tally the benefits of regular classes for each special education student but also calculate the risks. Educational harm can result when (a) decisions about students are based on stereotypes instead of individual strengths, (b) students are misplaced and left in placements that do not match their needs, or (c) students are included in regular classes without receiving services that comport with their IEPs. The word *harm* is mentioned only once in the IDEA, and that is in the least restrictive placement requirements.

Determining the Appropriate Least Restrictive Placement

To meet the standard of being individually appropriate, placement decisions in education and public policy are required to be made on a case-by-case basis using person-specific data, not philosophical arguments about social integration. In addressing the integration principle of the IDEA, educators and legal analysts have created analytic frameworks to ensure full consideration of the LRE placement requirements. One framework guides decisions about placement by using a sequential format that begins by considering whether the appropriate educational services written in the IEP can be delivered in the regular class if modified through the use of supplementary aids and services. If the answer is yes, then the regular class is the primary placement. If not, the decision-making team would move along the continuum of alternative placements one step at a time, from regular class to resource room, to separate class, to separate school, residential setting, hospital, or home, considering whether the appropriate educational services might be delivered with appropriate supports until the answer "yes" is obtained.

Another approach follows a set of questions to guide special education teams in making least restrictive placement decisions. The following questions, developed by legal scholar Mitchell Yell, address the components embodied in the LRE frameworks and rely on student-centered data collected throughout the LRE determination process:

1. Has the school taken steps to maintain the child in the general education class? What supplementary aides and services were used? What interventions were attempted?

2. What are the benefits of placement in a general education setting with supplementary aids and services versus the benefits of placement in a special education setting? What are the academic benefits? What are the nonacademic benefits (e.g., social communication and interactions)?

3. What are the effects on the education of other students? If the student is disruptive, is the education of other students adversely affected? Does the student require an inordinate amount of attention from the teacher and, as a result, adversely affect the education of others?

4. If a student requires a separate setting, are appropriate opportunities for integration available? In what academic settings is the student integrated with nondisabled students? In what nonacademic settings is the child integrated with nondisabled students?

5. Is the full continuum of alternative services made available across the system from which to choose an appropriate placement?

Although the IDEA prefers students to be included in regular classes with nondisabled students, the law does not require social inclusion. The courts, however, give careful scrutiny to decisions that place students in more restrictive settings.

Decisions regarding the least restrictive placement for any individual in public health, education, and human services require thoughtful analysis. Public policies acknowledge a rational basis for determining, on a case-by-case basis, that a particular person's appropriate treatment might be provided in specialized settings using differing strategies and more intensive supports.

Jean B. Crockett

See also Inclusion; Individualized Education Program; Individuals with Disabilities Education Act; Special Education

Further Readings

Crockett, J. B. (2005). IEPs, Least Restrictive Environment, and placement. In K. Lane (Ed.), *The principal's legal handbook*. Dayton, OH: Education Law Association.

Crockett, J. B., & Kauffman, J. M. (1999). *The Least Restrictive Environment: Its origins and interpretations in special education*. Mahwah, NJ: Lawrence Erlbaum.

Silverstein, R. (2000). The disability policy framework: A guidepost for analyzing public policy. *Iowa Law Review*, *85*, 1757.

Turnbull, H. R., Stowe, M. J., & Huerta, N. E. (2007). *Free appropriate public education: The law and children with disabilities*. Denver, CO: Love.

Yell, M. L. (2006). *The law and special education*. Upper Saddle River, NJ: Pearson.

LIFELONG LEARNING

Lifelong learning is an education trend that describes the engagement of adults in multiple educational opportunities across the life span. The learning

opportunities can be formal or informal, but these adult learning practices are distinguished from childhood learning because they are intentional, voluntary, and under the control of the adult learner. In today's society, adults may choose to explore further educational opportunities for vocational and personal enrichment or may be driven to more learning in the face of vast economic competition. However, the pattern of embracing multiple learning opportunities throughout adult life is the key factor in defining the lifelong learner. The choice to learn continuously across the life span demonstrates how some adults have adapted to shifting professional and personal expectations. With more adults seeking learning opportunities, colleges and universities have restructured their institutions to accommodate the lifelong learner. Similarly, human resource divisions within companies have been called upon to meet the needs of their employees who seek supplemental training or instruction. Lifelong learning, therefore, has left an imprint on the way adults think, work, and learn. The technology and information revolution of the past 10 years has amplified adult education opportunities, making it possible for more adults to have access to learning. In the future, educational psychology will need to examine the effects of continuous learning on adult cognitive development, the process of adult education, and the adjustment of adults who embrace or reject multiple learning opportunities throughout their lives.

Adult Learners in Higher Education

Colleges and universities have traditionally provided a formative experience for young people. But as adults began seeking additional learning opportunities in the past 30 years, the landscape of higher education has changed. Evening and weekend courses are now integrated in every college and university across the country, due to the demand for flexible course scheduling from adults who balance family and career obligations. Instructors and administrators now distinguish traditional-age students from nontraditional students and recognize their different needs, different expectations, and different ways of learning. Colleges and universities have begun accelerated degree programs for adults who began college courses earlier in life, left education to enter the workforce, and now are returning to classrooms with a rich set of life and work experiences. Accelerated degree programs often give these returning learners college credit for the acquisition

of those life experiences, with the twofold purpose of validating the experiences as learning and reducing the time it takes to obtain a degree. As a result of the increased demand for adult education, higher education has changed class schedules, courses, and degree programs to accommodate the new consumer, that is, the adult learner.

As adults have demanded more flexibility and choice in their educational experiences, technology has begun to play an exceedingly important role in how education is delivered. Colleges and universities now offer courses in their entirety over the Internet, complete with online class discussions, requirements for "posting" comments (the equivalent of class participation), and submitting all assignments electronically. In this new education delivery system, in lieu of having face-to-face classroom interactions, adults and instructors communicate via e-mail or chat rooms.

The virtual classroom meets the flexibility demands of adult learners, but educators speculate whether "e-learning" (i.e., electronic learning) achieves the traditional goals of higher education. To address this concern, some colleges and universities who offer online degrees have incorporated one or more physical face-to-face meetings between students and instructors. But adult learners, specifically lifelong learners, seek out higher education for a variety of personal objectives, some of which may not include the traditional objectives of higher education: to acquire critical thinking and writing skills, to solidify basic knowledge, and to be introduced to a particular professional field. As adults seek continuing educational experiences that support their lifelong learning goals, colleges and universities will need to adjust, again, their expectations and definitions of learning for these students. An instructor may find it difficult to expect the same motivation and quality of work from a traditional-age college student as he or she does from an adult learner who already has one degree but who enrolls in courses to learn more.

Adult Development Across the Life Span

As a movement, lifelong learning has changed the way educators and psychologists think about cognitive development. As adults continue to learn throughout their life span, previously held views of cognitive stage development may need alteration. Jean Piaget's theory of cognitive development holds that children

and adolescents move through stages of development based on their ongoing maturity combined with the influence of formal education opportunities. Specifically, Piaget theorized that adolescents reach formal operational thought, the highest stage of his cognitive development theory, only when they are afforded the experience of formal education. Erik Erikson posited that humans go through broad stages, not specifically cognitive, across the life span, which reflect traditional and historical biographies: establishing families, raising children, growing older, and evaluating one's life. But neither Erikson's maturation through the life span nor Piaget's cognitive development stages explain the developmental process of adults who continually seek out and are transformed by learning experiences.

Theorists such as William Perry suggested that the process of learning in college affects students cognitively and ethically in four broad categories. In the first category, dualism, Perry suggests that students focus on the extremes of intellectual ideas, that knowledge is right or wrong, good or bad, and that experts or authority figures hold the answers. In the next category, multiplicity, students move slightly from the dualist perspective in that they begin to let go of absolute thinking. This is a confusing time for students because they have to let go of some of their past ideology to make room for more knowledge, but in part they still seek to know the "right" answer. The third category, relativism, occurs when students resolve their "need to know the right answer" by framing ideas or choices as relative to the circumstances. Authority figures are seen as a source of information but not the only source, for in this stage, students begin to see their own ability to analyze and interpret information as significant. In the last category, commitment to relativism, students adopt the thinking that most ideas can be relative to the situation but that they must choose the best course of action or thinking for themselves.

Perry's stages of cognitive development were developed as a model of understanding how formal education affects the thinking of students. But the effects of formal education on adults who continue to pursue educational opportunities across their life span are not known. Drawing from Perry's work, it is possible to see that if formal education can facilitate cognitive development, then continuous learning across the life span would also have an effect on cognitive development. Jack Mezirow suggests that cognitive

development occurs when adults critically reflect on assumptions. For Mezirow, critically reflecting on assumptions requires examining suppositions of ideas and is the key factor in adult learning. As adults continue to pursue education across the life span, educational psychologists will have more opportunities to study how lifelong learning affects adult development.

Within the new field of *andragogy*, the study of how adults learn, researchers attempt to determine what strategies are best suited to help adults learn. Serguey Zmeyov says that adult learning is based largely on the learner's life experience, with all learning being filtered through those experiences where the "self" becomes an active mediator in the learning process. In this view, the learner and teacher act cooperatively because the adult learner has self-awareness and is actively involved in his or her own learning. But further research is needed to determine how choosing continuous learning as an adaptive function of postmodern life affects adult cognitive development.

Lifelong Learning in Recent Historical Context

In the United States, both the industrial revolution and the technology revolution solidified the global postmodern perspective by changing what and how adults learn. Folke Glastra, Barry Hake, and Petra Schedler have theorized that globalization and individualization have intersected to create societies that must continually adjust to stay current in hypercompetitive economies. These cultural and economic shifts are imperative to the understanding of how lifelong learning developed and how the pattern of lifelong learning represents a change in the way adults view and adapt to the world.

Unparalleled changes in the world of work during the past century set the stage for a new shift in thinking about the world of work. Specifically, modern workers transitioned from trades to factory positions during the Industrial Revolution. The factory worker learned to operate new machines or risk the unemployment line. The invention of the light bulb gave way to round-the-clock shift work, forcing workers to adapt to new living patterns. The transition from trade to factory work created a need for workers to embrace new machines, new techniques, and new learning about the world of work. In the spirit of enterprise and innovation, with workers being "trained" for specific jobs and not lifetime trades, the Industrial

Revolution moved forward, carrying many countries into economic prosperity.

World War II marked another major shift in the movement toward lifelong learning. The G.I. Bill launched hundreds of thousands of adult learners into the institutes of higher education. Eager not to disrupt the current economy with the sudden influx of returning soldiers, the government could not have foreseen the far-reaching effects of adult learners in classrooms. At that time, most Americans did not receive a college education, as it was preserved for specific professions and the financial aristocracy. With postsecondary education now financially accessible, American soldiers took advantage of this opportunity to integrate new knowledge into their lives with formal college degrees. This new generation of adult learners, now armed with degrees, launched into the booming postwar economy to eventually reshape the world of work. The social repercussions, however, could not have been easy to predict. Veteran fathers who returned home from the war and earned degrees went on to create families where higher education was an expectation for their children, not just an option. For the first time, middle-class children were being sent to college. In short, learning was becoming a social value.

Since the 1960s, cultural and social traditions have become decentralized from the human experience. This has had profound implications on the way adults work, learn, and make decisions about working and learning. Whereas social institutions reigned supreme in the mid-20th century, today, the individual is honored with the authority to choose the path that is best for him- or herself. Paul Belanger described the change as a new concern for personal growth and search for self-identity that signaled major changes in life patterns and increased life path diversification. This focus on the individual has created space for individuals to pursue more opportunities, more learning, and new identities.

Globalization and the Future of Learning

Today, as companies compete globally for business, employees must compete with more and more candidates for fewer jobs. Thus, the need to stay competitive, to learn more, and to produce more seems to be increasing. Today's employee is called upon to multitask, be multiskilled, and transfer those skills from organization to organization. Remaining competitive in the world of work means embracing new learning, new ideas, and new strategies. Adaptability is the heart of this new economic revolution, and lifelong learning represents the personal adaptation that more and more adults are choosing.

The topic of lifelong learning is currently being researched primarily outside of the United States, particularly by several European nations. The United Kingdom has a government-sponsored department of lifelong learning charged with supporting citizens in their vocational preparation and field transitions. As the international community examines the effects of globalization on the world of work, lifelong learning will continue to be explored. Formal and informal education demands are on the rise in every competitive industrialized nation. Lifelong learning is, therefore, a global education trend that demands further examination.

Tiffany Lynnette Arrington and Patricia A. Lowe

See also Adult Learning; Cognitive Development and School Readiness

Further Readings

Belanger, P. (1994). Lifelong learning: The dialectics of "lifelong educations." *International Review of Education*, *40*, 353–381.

Crow, S. R. (2006). What motivates a lifelong learner? *School Libraries Worldwide*, *12*, 22–34.

Glastra, F. J., Hake, B. J., & Schelder, P. E. (2004). Lifelong learning as traditional learning. *Adult Education Quarterly*, *54*, 291–307.

Gorard, S., & Selwyn, N. (2005). What makes a lifelong learner? *Teachers College Record*, *107*, 1193–1216.

Kett, J. F. (1994). *The pursuit of knowledge under difficulties: From self-improvement to adult education in America, 1750–1990.* Stanford, CA: Stanford University Press.

Love, P. (1999). Understanding and applying cognitive development theory. *New Directions for Student Services*, *88*, 5–99.

Mezirow, J. (1998). On critical reflection. *Adult Education Quarterly*, *48*, 185–198.

Zmeyov, S. I. (1998). Andragogy: Origins, developments and trends. *International Review of Education*, *44*, 103–108.

LITERACY

Four senses of the term *literacy* can be distinguished. First, literacy is simply the ability to read and write; it is the set of basic skills that provide individuals with

entry into the realms of written language and communication. Second, literacy is sufficient competence in reading, writing, and computation (numeracy) required to meet the demands of daily life. In this second sense, literacy represents a threshold of functional competencies that allow an individual to adapt to the demands of a complex and technologically sophisticated society. It is a pragmatic perspective linked to skills essential to employment, personal finances, consumer behaviors, and family life.

Third, literacy is the state of being highly educated, well read, well informed, and knowledgeable in the realms of culture, art, music, literature, history, economics, science, technology, mathematics, geography, politics, and public affairs—as such, it enables full participation in the cultural, economic, and political life of the society. A college-level education and preparation to enter a profession are commonly recognized as indications of literacy in this sense of the term.

Fourth, literacy is a characteristic of a social or cultural group. Literacy, in this sense, like language, can be regarded as varieties of cultural practices that entail the creation and use of texts. Individuals are literate to the extent that they may participate in the literate social activities that are important to the social groups to which they belong.

In education and psychology, theorists and practitioners have, for the most part, addressed literacy as the acquisition of the skills and competencies required for reading and writing. Throughout the past century, research and practice have been directed toward discovering the complex perceptual processes, cognitive systems, motivational factors, and social contexts that underlie an individual's ability to read and write and toward devising the most efficient methods to teach those skills. Basic research has drawn from all branches of learning theory, educational psychology, child and adolescent development, cognitive science, neurology, sociology, and cultural anthropology. Applied research has examined alternatives in teaching and instructional practices, in the development and implementation of curriculum and curriculum standards, in support for children and adolescents with learning disabilities and other special needs, in support of English language learners, and in assessment and evaluation practices.

Policy Debates

That literacy is essential for citizenship and for full participation in the cultural, political, and economic life of a free society is one of the most fundamental tenets of American education. There is agreement on all sides that illiteracy and marginal literacy preclude equal opportunities for employment, education, and civic engagement. Yet, despite decades of research and the repeated promises of developers of innovative approaches to teaching and learning, there are, by various estimates, between 20 and 50 million Americans who are functionally illiterate. The "literacy crisis" has prompted fierce controversies in educational theory and practice, which range from concerns for the relative success of alternative methods of instruction in reading and writing, to the value of liberal education, to the wholesale critique of public schooling in this society.

In recent decades, scholarly discourse on literacy has been replaced often by bitter debates and partisan political battles. In the 1960s, it was the "great debate"; more recently, it has been the "reading wars." The No Child Left Behind Act of 2001 codified elements of these debates into federal law, which requires that approved curricula for the elementary grades must now be based on these five essential components of reading instruction: phonemic awareness, phonics, vocabulary, fluency, and comprehension.

In an effort to address the public controversies surrounding literacy in the United States, a series of reports and policy statements have appeared from the national professional organizations: International Reading Association, National Council of Teachers of English, National Association for the Education of Young Children, the National Institute for Child Health and Human Development, the National Research Council, the National Reading Panel, the U.S. Department of Education, and many other agencies, associations, and organizations. The collective purpose of these efforts has been to establish clear understandings of what the best research has revealed about literacy and to outline areas requiring further research. Although there are continuing controversies about specific topics, such as the most effective approaches to phonics (or the learning of phoneme-grapheme, or letter-sound, correspondences), there is general agreement among psychologists and educators about how children acquire language and develop competencies in reading and writing, as well as other dimensions of literacy. The present consensus provides a sensible framework for understanding literacy development and the roles of parents and teachers who support children's literacy learning. In the

corresponding policy debates, there is also growing consensus, although some sharp disagreements remain concerning accountability, funding, and the allocation of resources to low income, linguistic minority, and special needs populations.

Evidence indicates that literacy learning is not in precipitous decline. Over the past 25 years, the National Assessment of Educational Progress has found that reading skills have remained virtually unchanged. These results suggest that schools continue to foster literacy development under conditions of increased immigration, shifting demographics, and the persistent achievement gap. American students continue to perform well in international comparisons in the fourth grade and at age 15, but relative performance declines sharply for the lowest quartile of U.S. scores in comparison to other countries. That suggests greater variability in performance and diminishing success for a significant number of lower-performing students as they enter the high school years. These data have prompted researchers and policy makers to continue the search for large-scale solutions to a perceived literacy crisis.

Literate Social Practices

Conceptions of literacy that emphasize individual skills and achievements may tend to overlook the essentially social nature of literacy. Literacy is more than simply the ability of individuals to read and write. It is a characteristic of social and cultural groups that engage in social activities that entail the creation and use of texts. Individuals are literate to the extent that they do participate in the literate social activities that are important to the social groups to which they belong. In this sense, children become literate by a combination of participation in informal engagements with texts (e.g., observing parents' literate activities, sharing in storybook reading) and in formal instruction with adults directed toward learning-specific aspects of the complex competencies that comprise literacy (e.g., deliberate teaching of letter-sound correspondences, reading comprehension strategies, procedures for organizing and composing an essay). What educators mean by literacy and how they understand the nature of learning and development will determine the nature of teaching practices that would create the conditions appropriate to invite learners to become competent participants in the literate social practices that have genuine meaning for them.

Literacy is participation in the socially embedded, purposeful activities that occur among members of a linguistic community (speakers, readers, and writers) and the texts they create and use to construct and share meanings and intentions. It follows that the development of literacy is the acquisition of the competence to fully participate in the literate social practices valued by that social group.

Children acquire cultural competencies by immersion and participation in the daily activities of the community. When oral and written texts (e.g., prayers and storybooks) are involved, children also become literate as they learn to participate in the use of these texts.

Upon entering school, children are exposed to a new set of social practices, a new community having its own rules concerning the use of language and literacy. Because all social practices are learned in relation to those already learned, the acquisition of the literate social practices of schools will depend on the cultural relationship between home and school.

Academic English

Research and practice that focuses on English language learners, and on speakers of dialects that diverge significantly from standard English, have drawn attention to the distinctive characteristics of academic English and its relationship to literacy. It is widely recognized that students who complete high school and are well prepared to enter college need to have command of academic English, which includes rich vocabulary, advanced reading comprehension, control of complex syntax, and competence in standard written English. It is also widely recognized that schooling alone cannot provide sufficient opportunities to achieve all the necessary competencies. Regarding vocabulary, for example, it is estimated that, in one school year, it might be possible to directly teach about 1,000 new words, but it is also estimated that to be well prepared to enter college requires a receptive vocabulary of 80,000 words. If direct teaching is insufficient to accomplish this goal, then informal and incidental learning are required. Moreover, the student must have access to Latinate words (e.g., *inundate* for *flood*) and other learned borrowings into English.

Access to academic English requires extensive reading of complex texts and routine exposure to those literate social practices that typify literacy in the

sense of becoming highly educated. Students whose parents are well educated will tend to have greater exposure to academic English. For all other students, the schools become the chief source of opportunities to acquire the discourse patterns used in higher education and the professions.

The core of everyday English, as is well known, comes down to present times from Anglo-Saxon (as well as Old Norse—examples include *get, want, give*). Yet, it is estimated that perhaps 75% of the English lexicon derives from French, Latin, Greek, and assorted other languages, and many of these are regarded as scholarly or academic words not generally used in everyday spoken communication in the home and community. Access to literate language in the sense of academic English is also a condition of access to literate social practices that provide entry to higher education and corresponding economic opportunities.

Emergent Literacy

Research efforts of the past 40 years have made tremendous progress in understanding the genetic, neurological, physiological, cognitive, and social systems that make possible the acquisition of language and social communication. Among the core findings are the following:

1. All human societies use language, and all children acquire their native language—with the relatively rare exception of those who suffer significant impairments in one or more of the systems that support language.

2. The 5,000 to 6,000 languages that exist among the cultures of the world today have abstract features in common, and all particular languages are but variants of "human language"; the common or universal features of human languages are given by human genetic endowment.

3. Native language acquisition occurs rapidly without formal or expert instruction. By age 5 or 6, children have acquired the core features of all the rule systems that govern language form and use.

4. The broad stages of language acquisition are invariant across cultures and languages, although at finer levels of analysis, patterns of individual differences emerge.

5. Language is acquired in use; it is a meaning-making process through which children construct

or reinvent the language while engaged in genuine communication.

6. Specific features of the social environment account for the specific language children will acquire; that is, children will acquire the linguistic patterns of their parents and members of their immediate communities.

7. Specific features of the linguistic environment account for some of the differences in the nature of language acquisition between and within groups, for example, using specific discourse styles, developing a large vocabulary, or learning modes of politeness.

Human language is a complex system that relates sound to meaning. An adequate description of language requires, at minimum, the recognition of multiple, interacting subsystems, each of which must be acquired: pragmatics, semantics, syntax, morphology, phonology, and the lexicon. Language production and comprehension require the simultaneous, rapid processing of information from all subsystems, which are acquired without conscious reflection or formal instruction.

Literacy develops upon the foundation provided by language acquisition for children who are raised in a literate social environment. From the earliest recognition of street signs and product logos, to the earliest experiences with books, stories, and other texts, children's literacy development interacts with language development. For example, lexical development (vocabulary) is an astoundingly rapid process. By age 3, a typical child will have a receptive vocabulary of 1,000 words. By age 6, a conservative estimate will be 10,000 words. Children who are richly exposed to the language of books will tend to acquire a more varied, rich vocabulary, which will, in turn, enable them to successfully engage with ever more challenging texts.

Children who are exposed to the alphabet from an early age will begin to associate letter names with the visual letters and later will begin to associate some speech sounds with letters. Nursery rhymes and children's songs and poetry will help children recognize (and become "aware of") features of speech sounds and words, for example, rhymes linking ending sounds of words and alliteration drawing attention to beginning sounds. Taken together, such experiences support phonological development and, in particular, phonemic awareness.

Family literacy provides children with many opportunities for learning the basic features of oral and written language. When parents read to children, or tell stories, sing songs, share family histories, say prayers, and so forth, they are providing opportunities for children to experience the pleasure of powerful texts and to become sensitive to the structures of narrative. Children's listening comprehension for texts will far exceed their independent reading skills for most of the elementary school years.

Finally, children who are given opportunities for writing in the preschool years will have further occasions for linking letters to sounds to meanings. Whether with magnetic letters, alphabet blocks, pens and pencils, or computer keyboards, children will string together letters in an attempt to "write" and to have adults "read" their creations. What children begin to learn is that written symbols correspond to speech. All early literacy experiences provide children with richly important opportunities to learn the relation between print and language.

Under the circumstances of exposure to a genuinely bilingual home or community, children from infancy can develop two languages simultaneously. In many places in the world, bilingualism is the norm. When children are immersed in a second language during the preschool years, and where that language is used for the full range of communicative purposes, children can achieve fluent bilingualism to the same levels of competence of monolingual children within a few years. However, when children first encounter a second language in the context of an elementary school classroom, where instruction (and especially literacy instruction) is primarily in the second language, then the children will face greater challenges in acquiring that language and developing literacy. Greater home support for the two languages and literacy in both languages will make it more likely for children to succeed.

Children whose home language is other than English may appear to be fluent in English because they have acquired the "language of the playground" during preschool and kindergarten experiences. However, some of these children will exhibit fluency but not have achieved proficiency in the language to the level that will readily support literacy development. A first grader whose home language is other than English may appear fluent even if the child has command of only 1,500 words of English. That number, however, is far less than the 10,000 or more words known by the monolingual English child growing up within a literate family. On the other hand, the child whose home language is, for example, Spanish, and whose parents provide a literate home environment in Spanish, will acquire the essential early literacy competencies in Spanish and be better prepared to transfer those competencies to the English-speaking context.

Beginning Literacy

Children who enter school with those experiences that support language development and early literacy learning begin more formal literacy instruction with

- knowledge of the language,
- an extensive vocabulary,
- initial phonemic awareness,
- knowledge of environmental print,
- knowledge of the alphabet,
- awareness of the organization of books and stories,
- experience with writing, and
- knowledge of some letter-sound correspondences.

Some children will also know some words and names by sight. They may also have memorized favorite stories or songs which they can "pretend" to read, and they may have dictated or written their own stories with the help of parents, siblings, or other caregivers. A few children will have become "early readers" who can read independently (whether silently or aloud) from so-called chapter books that are commonly read in third or fourth grade.

Children who have *not* been exposed to the informal opportunities for literacy learning in the home will enter kindergarten and first grade with a distinct disadvantage. In addition, children whose home language is other than the language of the schools will have the greatest challenge if they are expected to both learn the language of the school and begin to read and write for the first time in the unfamiliar language.

Kindergarten and first-grade classrooms should provide a literate environment; that is, teachers and students should engage in literate social practices for authentic purposes in order to continue the powerful informal learning that underlies language acquisition and home literacy learning. Children need to share with each other and to participate in reading and writing activities with competent adults who can model their own literacy. Teachers need to read aloud from the best of literature for children, as well as encourage

children to read for themselves by first engaging in shared reading and guided reading. Similarly, teachers need to support children's writing with varied opportunities, ranging from taking dictation, to shared and guided writing, to independent writing.

Informal learning should be complemented by formal instruction. To succeed in becoming literate, some children may need only to be exposed to competent models. Most children, however, will also require direct instruction that will guide them through the intricacies of letter-sound correspondences, the vast number of irregularities of English spellings (both for reading and writing), the complexities of comprehension across diverse genres, and the challenges of writing narrative and exposition for various audiences. A number of children—estimates range from 1% or 3% all the way to 10% or 20%—will experience greater difficulties achieving literacy, even when English is their first language and their parents provide a literate home environment. These children will require further, expert direct instruction, such as that provided by Reading Recovery, or other one-to-one tutor programs. Research clearly indicates that early, intensive intervention is more likely to be successful than later interventions.

Direct, planned, explicit instruction may take many forms. Researchers and practitioners continue to dispute the exact nature of direct instruction. Some call for rigid scripts, lesson sequences, highly controlled texts, and heightened attention to isolated, specific features of print, for example, letter-sound correspondences (sometimes referred to as phoneme-grapheme relations). Others call for direct instruction in the context carefully chosen, quality literature that will maintain a meaningful context while also allowing success (reading with comprehension) and attention to specific features of print. Collections of recommended books are sometimes "leveled" in an effort to more accurately determine appropriate "instructional-level" materials. Greater training of teachers and greater varieties of books and materials are required for the latter approach.

The general consensus of teachers in the field is that a balanced or integrative approach is best, with flexibility in arrangements of whole class, small group, and individual instruction, which itself ranges from well-crafted informal learning environments to expertly planned formal learning opportunities. The primary goal of an integrative approach is to sustain meaningfulness within an authentic literate classroom environment, while also providing explicit, direct instruction where it is judged to be most needed, including judicious attention to practice with isolated features of the language. Children need time to read and write for pleasure and for the range of authentic purposes. Instruction should build on prior knowledge, provide support where needed (scaffolds), and present challenges in reasonable measure to advance children's competencies. Moreover, instruction should be guided by careful observation and ongoing assessment of children's successes and struggles; thus, instruction is responsive to the requirements of specific children in specific circumstances.

Reading and writing, like language processing (listening and speaking), require the simultaneous, rapid processing of multiple levels of information; this is also called parallel processing. At a minimum, the levels include

- orthographic (letters, spellings, capitalization, punctuation),
- phonological (sounds),
- morphological (word forms, plurals, tenses, prefixes, suffixes, etc.),
- syntactic (clause structures and sentences, larger discourse),
- lexical (properties of words, including sounds, spellings, and meanings),
- semantic (meaning and comprehension),
- pragmatic (intentions, e.g., narrative and expository), and
- cultural (e.g., assumptions, background knowledge, and cultural expectations).

Fluent, independent reading and writing require competence within each of these levels of processing. Instructional approaches that substantially neglect certain levels while attending exclusively to others, or that focus on reading while neglecting writing, will be less successful in supporting full literacy development.

Development
Toward Independence

If English provided a perfect alphabet system, then it might be possible for children to learn the alphabetic principle (letter-sound correspondences) in 2 or 3 months in first grade, and then all other language arts instruction would be directed toward developing language: greater vocabulary, more subtle comprehension, writing for an audience, greater knowledge of

the physical and social world through texts. Learning to read, as such, would be very easy. The complexity of the English orthography tends, however, to require 3 or more years for children to gain independence in reading sufficient to take full advantage of their oral language competence. Learning to spell in English becomes a lifelong challenge for many, including many well-educated adults.

Learning to read (accurately, fluently, and with comprehension) requires both frequent access to demonstrations of mature literacy by competent models and continued direct instruction in effective strategies throughout elementary school. The same may be said for learning to write (including spelling, punctuation, and conventions of composition). Both reading and writing should be blended with other aspects of language development and language arts instruction. Discussions about the role of phonics in reading instruction often ignore that in learning to write, children are confronted with the complementary problem of discovering the appropriate letters to represent sounds. They soon learn that letter-sound correspondences can take one only so far; word knowledge will be required to make further progress (e.g., consider homonyms, such as *pear, pair*, and *pare*, or words with a common root, such as *native, nation*, and *nature*).

Thus, second- and third-grade instruction should continue the practices initiated in kindergarten and first grade, carefully advancing the richness and the challenges of the literate environment (sustaining informal learning), while continuing direct instruction where needed as children encounter ever more complex features of English. Exposure to larger vocabularies, wider varieties of text, and more complex ideas requires expert instruction that includes continuous assessment of the progress of individual children and the willingness to make instructional decisions based on that information. An integrative approach remains the consensus of experts in the field, though differences of opinion persist concerning the day-by-day, lesson-by-lesson recommendations. Those differences remain largely matters of emphasis.

Developing Proficiency and Fluency

Extending the practices outlined in the previous section, the goals of the intermediate grades shift from learning to read to reading to learn—across the curriculum. Reading and writing are directed toward developing knowledge and understanding in an ever-expanding range of texts, genres, and purposes. Students continue to need opportunities for collaborative, informal learning within an authentic literate environment where teachers continue to model their competencies. They also continue to need explicit demonstrations, direct instruction, opportunity to practice, encouragement to achieve fluency as well as precision, and high expectations. Students need to engage in literacy for authentic purposes, for example, to write for genuine audiences (beyond the classroom teacher alone) and to read for interest, pleasure, and self-directed learning.

Those students who, for various reasons, are still struggling with beginning reading and writing require special opportunities, individually or in small groups, for accelerated learning, guided by master teachers with the goal of having them catch up to their peers. Such instruction will be directed toward the students' greatest challenges (e.g., phonics, comprehension strategies, or gaining necessary knowledge of English) and will support and build upon those areas within which the students have the greatest successes. Late interventions are less likely to achieve rapid results than earlier interventions, and they will tend to require greater allocation of resources. Support for home literacy and English language learning should be incorporated where appropriate.

Literacy Programs

In the United States, a long tradition of cooperation exists between school systems and numerous publishers and curriculum developers, who also work closely with national professional associations and universities, in the design of literacy programs. At present, more than 20 well-known reading and writing programs are available from publishers, foundations, and universities, and each offers attention to

- mechanics of reading (graphophonemic details, visual processes),
- reading comprehension (strategies, purposes, and critical reading),
- literature (varied genres),
- spelling and punctuation,
- word knowledge, vocabulary,
- grammar and conventions of written language,
- writing expressively and persuasively (with revision and editing),

- approaches to maintaining motivation and pleasure in literacy,
- approaches to integrating the language arts and literacy across the curriculum, and
- using reading and writing as tools for thinking and learning about the world.

Each also provides some form of assessment or recordkeeping procedures to aid teachers and administrators in ongoing program planning. Most also include attention to new technologies. Some focus on children with specific difficulties (e.g., dyslexia or other reading disabilities), and others focus on the needs of children acquiring English as a second language. Yet, all draw from the same understandings of what reading and writing entail and what variety of classroom activities are available to support literacy development. Publishers and program developers have combined, balanced, and integrated the different elements in the previous list, each with some difference in emphasis.

There is no experimental research that definitively identifies one variant program as superior to another. Such research is not even possible because of the practical limits of experimentation with human subjects in the real world. When critics demand proof from randomized, experimental field studies, they must settle for small-scale studies completed with small populations in specific locations, with particular, uncontrolled local conditions. Sorting through this research was indeed the task of the National Research Council (see *Preventing Reading Difficulties in Young Children*, 1998) and the National Reading Panel (results published by the National Institute of Child Health and Human Development, in 2000).

Their findings mention a few specific programs but focus on underlying, common features that have been shown to contribute to the success of various programs.

Socioeconomic Conditions

All children throughout the world acquire their parents' language—the "mother tongue." All learn the discourse patterns of their families and communities. Language is universal, but literacy is not. In addition, the conditions of schooling and literacy are not equally distributed in society.

It is well established that socioeconomic conditions are powerful predictors of both literacy and school success. Where it is possible to establish good measures of

family literacy (levels of parents' education and literacy practices in the home) and sensitive measures of family support for schooling and literacy (e.g., parents' beliefs about the value of school success and literacy), these factors can be combined with economic conditions to produce multiple correlations that reach .8 or higher.

What this suggests is that the unique contributions of the schools (and teachers) should be understood in relation to other social conditions. It follows that research that focuses exclusively on school-based conditions will be overlooking important, powerful conditions. Also, it follows that only marginal differences will be found by new research into competing school-based programs, which, with minor differences, are themselves guided by the same body of knowledge and understandings of literacy development.

In the current debates on literacy, there is an astounding silence on these issues. Most reports focus on the school or the classroom and largely avoid the larger social issues. In the call for consensus and the pressure for "best practices," there is an explicit bias toward considering only the scientifically based research, which is defined to mean only rigorously controlled experimental and quasi-experimental designs. Because students (and parents) cannot be randomly assigned to socioeconomic categories, those variables, arguably the most powerful, are recognized in most research designs only as the "co-variates"—that is, efforts are made to control them rather than to directly investigate them. Other research approaches, especially ethnographic models, which can more fully investigate sociocultural and economic influences, are given less attention in the current reports. Quantification, control, and replication are preferred over long-term, careful documentation of complex systems.

Future Directions

The biological and social foundations of language acquisition are now reasonably well understood. In a similar sense, the nature of family literacy and school curricula that support reading and writing and the full scope of literacy development are well understood. Questions remain as to how best support language and literacy for children with special needs and for children who are learning English as a second language. There is consensus, however, that excellent and expert teaching succeeds when conditions are most conducive, especially in relation to community

support and to institutional variables such as the number of children who are to receive attention at the same time from the teacher. Children benefit the most when teachers can move flexibly among large-group, small-group, and individual modes of instruction.

The general consensus of teachers in the field is that a balanced or integrative approach is best, with both expertly planned formal learning opportunities and well-crafted informal learning environments (including competent models and demonstrations). The primary goal of an integrative approach is to sustain meaningfulness within an authentic literate classroom environment, while also providing explicit, direct instruction where it is judged to be most needed. Moreover, instruction should be guided by careful observation and ongoing assessment of children's successes and struggles; thus, instruction will be responsive to the requirements of individual children in specific circumstances.

Continued investigation of the social, linguistic, economic, and political conditions of literacy development should be included within the current attention to best practices and experimental research efforts to compare one specified program with another. Only then will we have the opportunity to move closer to universal literacy and full participation in the cultural, economic, and political life of this society.

William T. Stokes

See also Assessment; English as a Second Language; Reading Comprehension Strategies

Further Readings

Adams, M. J., (1990). *Beginning to read: Thinking and learning about print.* Cambridge: MIT Press.

Berliner, D., & Biddle, B. (1995). *The manufactured crisis: Myths, fraud, and the attack on American public schools.* Reading, MA: Addison-Wesley.

Chall, J. S. (1983). *Stages of reading development.* New York: McGraw-Hill.

Clay, M. M. (1991). *Becoming literate: The construction of inner control.* Portsmouth, NH: Heinemann.

Cunningham, P., & Allington, R. (2007). *Classrooms that work: They can all read and write* (4th ed.). Boston: Pearson.

Durkin, D. (1966). *Children who read early.* New York: Teachers College Press.

Holdaway, D. (1979). *The foundations of literacy.* Sydney, Australia: Ashton Scholastic.

Kozol, J. (1985). *Illiterate America.* New York: New American Library.

National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (NIH Publication No. 00–4769). Washington, DC: U.S. Government Printing Office.

National Research Council. (1997). *Improving schooling for language-minority children: A research agenda* (D. August & K. Hakuta, Eds.). Washington, DC: National Academy Press.

National Research Council. (1998). *Preventing reading difficulties in young children* (C. Snow, M. S. Burns, & P. Griffin, Eds.). Washington, DC: National Academy Press.

National Research Council. (1998). *Starting out right: A guide to promoting children's reading success.* Washington, DC: National Academy Press.

Pinker, S. (1994). *The language instinct: How the mind creates language.* New York: William Morrow.

Longitudinal Research

Longitudinal research is generally defined as studies that investigate change over time with the same cohort(s). This definition characterizes four key features of longitudinal research.

First, by *change*, it means that longitudinal studies focus on a dynamic process rather than a static status as the outcome variable. A process is a series of changes from the initial status to the final status through various intermediate statuses. In a methodological term, one status is often called one *wave*. Thus, longitudinal studies use multiple waves to represent a process.

Second, by *over time*, it means that longitudinal studies must include the time dimension as the fundamental predictor. A multiwave process can be considered as a function of time, and various substantive predictors (e.g., gender or IQ scores) can be used to explain the process, as in nonlongitudinal studies.

Third, by *with the same cohort(s)*, it means that longitudinal studies typically measure the same group of individuals (i.e., one cohort) repeatedly over time to examine the intraindividual changes (e.g., each student's reading development) as well as interindividual differences (e.g., gender differences in reading development). In some cases, a longitudinal study repeatedly measures just one same individual (e.g., the single-subject study) or multiple cohorts (e.g., the

cohort-sequential study). Longitudinal studies focus on changes within the same cohorts, whereas cross-sectional studies focus on differences across different cohorts.

Fourth, by *investigate*, it means that longitudinal studies are not limited to repeated measurement or analysis of change over time. Instead, longitudinal design, longitudinal measurement, and longitudinal analysis represent not only the three areas of longitudinal research methodology as an established field of study but also three major phases of a longitudinal study.

Learning, teaching, education, development, intervention, and various psychological processes are essentially intended to produce desirable changes among human beings. Thus, focusing on change over time, longitudinal research can be considered the method of choice for the social and behavioral sciences in general and educational psychology in particular. This entry presents basic methodological issues of longitudinal design, measurement, and analysis and provides real-life examples of empirical longitudinal studies.

Longitudinal Design

Longitudinal design concerns how to collect longitudinal data to address longitudinal research questions. To study children's social development, for instance, Anthony Pellegrini and his collaborators conducted a longitudinal study to examine games (e.g., chanting, chasing, playing balls, jumping rope, clapping) that the first graders played on the school playground. Specifically, they addressed two major research questions: whether there were any changes in games and why these changes occurred. They chose to use a single cohort, a total of 77 students in all first-grade classes from two neighboring schools. They collected three waves of data, mainly observing the change in frequency and variation of these students' games that were played in September, January, and May. They focused their observations on children's games at recess, a good time window to see children's free interactions at school, during their entire first-grade year as their first experience of full-day mandatory schooling in the 1995–1996 school year, with an interval of over 3 months. They found that there was a significant increase in frequency of game activities over the entire year, and this increase was significantly influenced by gender and ethical differences. This example illustrates that at least four specific

issues, the question design, cohort design, wave design, and time design, need to be considered carefully in designing a sound longitudinal study.

First, a longitudinal study must address research questions that are clearly related to change over time. There are generally three kinds of research questions with a longitudinal nature: (1) Does change occur over time? To answer this question, researchers must examine the initial status and final status of a process (e.g., examine the frequency of first-graders' games played on the school playground in September 1995 to May 1996). (2) How does change occur over time? To answer this question, researchers must look at the initial status, final status, and a few intermediate statuses of a process (e.g., observe the frequency of first-graders' school playground games in September, January, and May). (3) Why does change occur over time? To answer the *why* question, researchers need to examine not only a multiwave process but also various substantive predictors that might contribute to the process (e.g., estimate both gender differences and race differences in changes of school playground games in the first-grade year).

Second, it is important to decide how many cohorts will be used. In education and psychology, a typical longitudinal design is a quasi-experimental design; that is, one single nonrandomized cohort is investigated repeatedly over time. To address different research questions, longitudinal researchers have used a wide variety of longitudinal designs, such as intervention design, cohort-sequential design, microdevelopmental design, life-span study design, nonexperiment design, and even occasionally true-experimental design that randomly assigns human subjects to experiment and control groups to increase internal validity. Longitudinal methodologists Judith Singer and John Willett recommended the cohort-sequential design for effectively collecting longitudinal data in a limited time period by studying multiple age cohorts simultaneously.

Third, it is also important to determine how many waves of data need to be collected. Two waves of data enable the researcher to address only the research question of whether change occurs over time, whereas three or more waves of data help the researcher to address the question of how change occurs over time. Singer and Willett found that increasing the number of waves of data collection substantially reduces the standard error of individual rate of change and significantly increases the reliability of estimated rate of change and thus suggest collecting extra waves of data

even at the cost of the total sample size. Two types of special longitudinal designs, the time series design and single-subject design, normally include more than 10 waves of data to discover underlying patterns.

Finally, longitudinal researchers need to consider the issue of time thoughtfully, because the precision of estimating individual change depends more on the duration, frequency, interval, and timing of the waves of data collection. For instance, how long must the study last (e.g., 1 semester, 3 years, or 4 decades to capture the target behavioral change)? When will the study start (e.g., age 13 or age 17 to reduce the cohort effect)? What interval should be used (e.g., four times per semester or once per decade to reduce the practice effect)? What specific time is the best time to collect data (e.g., before the No Child Left Behind Act was implemented or after the Act to avoid the historical effect)? Recently, differing from the time design of conventional longitudinal studies of change over years or decades, microdevelopmental designs that intensely investigate over a short period of time have been extensively used by developmental psychologist to capture ongoing transitions in a critical period of time (e.g., examining second-graders' change in the use of different arithmetic strategies 10 times within 1 week).

Longitudinal Measurement

Longitudinal measurement concerns how to develop valid, reliable, and equitable measures for examining behavioral change over time. A longitudinal study normally involves three types of measures: (1) the outcome variable, which can be either continuous or dichotomous but must be time-variant to represent a multiwave process (e.g., initial, middle, and final IQ scores); (2) the time predictor, which can be either discrete or continuous but must be sensitive enough to capture change over a specific time period (e.g., studying early reading development for 3 years with the interval of 3 months); and (3) the substantive predictor, which can be either continuous or dichotomous and either time-variant or time-invariant (e.g., gender and change of socioeconomic status over 10 years). These three types of variables form the empirical foundation for subsequent longitudinal analysis of data, and whether these variables are valid, reliable, and equitable determines the quality of a longitudinal research.

Marlene Schommer and her collaborators, for example, examined whether and why there was improvement in high-school students' epistemological beliefs

(e.g., "really smart students don't have to work hard"). They developed a 63-item questionnaire on a 5-point Likert scale with adequate validity and reliability and used it to assess students' epistemological beliefs twice as the outcome variable. They focused on the 3-year period between the freshman year to the senior year in the high school between January 1992 and January 1995, with 3 years as the interval, as the time predictor. They used gender as the substantive predictor. With these measures, they found significant improvement in epistemological beliefs as well as gender main effect (girls had better beliefs) and gender-year interaction effect (gender difference was greater after 3 years). In another longitudinal study, David Kaplan and Sharon Walpole examined 3,575 students' early reading development across kindergarten and first grade. Their outcome variable was four-wave dichotomous measures of reading skills, a pass or fail score at a particular skill level on a Gutman scale. They focused on a 2-year period (1998–1999) at four time points (fall kindergarten, spring kindergarten, fall first grade, spring first grade). The substantive predictor was poverty status in 1998, a dichotomous predictor of whether the child's household was below or above the poverty line. Based on these measures, they found that children living below poverty are less likely to experience successful reading transitions than their above-poverty peers.

A unique and controversial feature in longitudinal measurement is that normally the same instrument ought to be used repeatedly over time to generate time-variant outcome variables or substantive predictors. This produces the test practice effect, or penal conditioning effect, which reduces validity and reliability of measurement, especially while many waves of data are collected and the between-wave interval is small. Some methodologists suggest avoiding change scores completely, whereas others have considered recently that this practice effect is small. In addition, some methodologists suggest developing equitable versions of the same instrument for repeated use (like ETS in building a large item pool for the TOEFL or the GRE), some encourage using dichotomous variables with clear-cut criteria, and others advocate unobtrusive measures through nonactive observations.

Longitudinal Analysis

Longitudinal analysis concerns how to analyze longitudinal data to address various kinds of research questions about change over time. Regression analysis and

repeated measures analysis of variance (ANOVA) are two, widely used, simple statistical methods for longitudinal data analysis in educational psychology research. For instance, to examine what factors influenced high-risk children's math achievement scores, Shane Jimerson and others conducted a regression analysis by using five predictors (students' years in special education, behavior assessment scores, parental involvement, home environment, and socioeconomic status) to predict differences in math scores between Grades 1 and 6 as the outcome variable. The results indicate that only socioeconomic status was the significant predictor of the math score difference. Here, regression analysis can handle only two-wave data when the outcome variable is either the final status or the difference between the initial and final status. It is the most rudimental data analysis method. In another study, Pellegrini and others ran repeated measures ANOVA to examine whether and why there were any changes in children's playground games at recess across the three time points. They used the frequency of the total games observed at three time points as the within-subject variable, and gender and race as the between-subject variable. They found significant effects of time, gender, and race. Better than regression analysis, repeated measures ANOVA can handle three or more waves of data by treating one multiple-wave outcome variable as multiple dependent variables in the multivariate analysis of variance (MANOVA) context.

Among various newly emerging advanced longitudinal data analysis methods, multilevel growth modeling and survival analysis are among the best tools to effectively analyze multiwave longitudinal data and address complex research questions. For instance, Zheng Yan fitted a basic two-level growth model to examine how and why a group of students' performance of using a statistic program changed across four waves for one semester. The level-1 model fitted individual trajectories of students' performance at four time points to describe within-individual change as a function of time, and the level-2 model further examined the intercept and slope of the average fitted lines to explain between-individual difference as a function of four substantive predictors. The findings indicate that only students' previous experience of using computer network systems affected the initial status of learning and that none of four substantive predictors influenced the rate of learning. The multilevel growth model is powerful, flexible, and integrative in analyzing complex longitudinal data that are not only multiwave but also multilayer (e.g., using structural equation modeling [SEM] for manifest and measure variables), multilevel (e.g., using hierarchical linear modeling [HLM]), multivariable (e.g., using MANOVA), and multipath (e.g., using path analysis). While the multilevel growth model can effectively analyze complete *continuous* longitudinal data, survival analysis can effectively analyze complex *dichotomous* longitudinal data. These advanced methods overcome the obvious weakness of regression analysis and repeated measures ANOVA and create new opportunities to entertain rich longitudinal data and address complex longitudinal research questions.

Future Directions

Although longitudinal research has been performed for more than a century, it has not yet reached its full potential. It is currently enjoying a fast growth and will become a true method of choice in the social and behavioral sciences. The promising directions of longitudinal research might include (a) addressing research questions about very complex and dynamic changes over time in the real world; (b) performing Internet-based automatic data collection in authentic settings and collecting natural data through the use of digital video technology; (c) developing computer-generated measurements and unobtrusive measurements; (d) conducting nonlinear dynamic modeling, computer simulation, and mixed qualitative-quantitative data analysis; and (e) witnessing productive dialogues between longitudinal methodologists and longitudinal practitioners.

Zheng Yan and Lailei Lou

See also Cross-Sectional Research; Measurement

Further Readings

Jimerson, S., Egeland, B., & Teo, A. (1999). A longitudinal study of achievement trajectories: Factors associated with change. *Journal of Educational Psychology*, *91*(1), 116–126.

Nesselroade, J. R., & Baltes, P. B. (1979). *Longitudinal research in the study of behavior and development*. New York: Academic Press.

Pellegrini, A. D., Kato, K., Blatchford, P., & Baines, E. (2002). A short-term longitudinal study of children's playground games across the first year of school: Implications for social competence and adjustment to

school. *American Educational Research Journal*, *39*(4), 991–1015.

Schommer, M., Calvert, C., Gariglietti, G., & Bajaj, A. (1997). The development of epistemological beliefs among secondary students: A longitudinal study. *Journal of Educational Psychology*, *89*(1), 37–40.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York: Oxford University Press.

Willett, J. B., Singer, J. D., & Martin, N. C. (1998). The design and analysis of longitudinal studies of development and psychopathology in context: Statistical models and methodological recommendations. *Development and Psychopathology*, *10*(2), 395–426.

Yan, Z. (2006). Different experiences, different effects: A longitudinal study of learning a computer program in a network environment. *Computers in Human Behavior*, *22*, 364–380.

Long-Term Memory

Long-term memory (LTM) refers to people's vast storehouse of retrievable information other than perceptual and short-term memory. It usually remains dormant until activated by a particular stimulus event and is divided into a couple of components: *episodic LTM*, which contains individuals' personal histories, their recollections of what, when, and where events have occurred in their past; and s*emantic LTM*, which is individuals' storehouse of knowledge that is not time dependent. Research of episodic LTM focuses on two main categories: forgetting and remembering.

Theories of Long-Term Memory

An early theory of LTM was the perceptual moment hypothesis. This theory proposed that people encode information in 100-millisecond bundles and store them into memory as an unbroken record, a continuous loop—10 snapshots per second, 600 per minute, and more than 500,000 per 16-hour day. This theory did not explain how or where these memories were stored or how they could be retrieved. Presumably most memories remained hidden from people except for unexpected remembrances or purposeful efforts to recall. Dreams, often comprising mindless yet vivid segments, might draw their substance from random samplings from this loop.

Subsequent research by the eminent neurologist Wilder Penfield lent credence to this theory. Penfield operated on patients who were afflicted with violent epileptic seizures. Removal of damaged regions of the brain often reduced the severity of the seizure. Before surgery, and with the patient fully conscious, Penfield applied a mild electrical current (which produces no pain) to exposed regions of the brain. How the patient responded to the current helped Penfield distinguish damaged tissue from healthy tissue. Many points of stimulation produced utterances and body movements. However, when regions of the temporal lobe were stimulated, vivid memories moving forward in real time were reported, as if this continuous "memory loop" had been engaged at some temporal location. One patient, a stenographer 10 years earlier, stated, "I could see the desks. I was there, and someone was calling to me, a man leaning on a desk with a pencil in his hand." Another reported, "I hear music again," and began to hum along, later noting, "There were instruments. It was as though it were being played by an orchestra. Definitely it was not as though I were imagining the tune to myself. I actually heard it." Penfield, who operated on more than 1,000 patients, concluded that he had tapped into a "stream of consciousness" reflecting memories stored many years earlier.

However, not all patients reported re-activated memories, and other scientists could not replicate Penfield's findings. Other scientists simply were skeptical of the claims by Penfield, and this line of research largely disappeared.

Today, researchers know that some of Penfield's speculations, provocative as they were, must be incomplete if not wrong. For one, memories cannot be a veridical encoding of an event, like an unbiased camera recording of each day's activities. Rather, memories often contain embellishments, added at later points in time. Memories can even be created for an event that never occurred at all. In short, people store not only what they have experienced but also significant "add-ons" that can alter or distort the original memory. Second, it seems likely that people filter their experiences based on what engages their attention and emotions. What escapes their attention is likely not encoded at all and, therefore, is absent from memory.

To appreciate the complexities of human memory, it is important to start from a somewhat different but broader perspective.

Multiple Memories

People have more than one kind of memory. They have *perceptual memories*, one for each modality, that last literally for a few fleeting milliseconds. If you are careful, you can catch these brief memories—watch something involving motion and quickly close your eyes. You may capture a vivid image of the event before it fades. These brief sensory stores are probably the entry points of most memories.

People have a *working* or *short-term memory* (STM) as well, and this is where their consciousness resides. Working memory is limited in what a person can hold or attend to—a handful of digits or words—and today's study of working memory has linked it to the comprehension of ongoing events, as in reading of a book, following a conversation, or solving a puzzle. In this ever-changing world, these dynamic events enter and leave people's consciousness, leaving behind a trace that lasts no more than 15–30 seconds.

People also have LTM, described in the opening paragraph to this entry as consisting of episodic memory and semantic memory. A person may forget what he or she had for lunch last Monday (episodic memory) but not the meaning of lunch itself (semantic memory). LTM also includes *procedural memory*, which is people's knowledge of sequences, events that occur in temporal order, such as riding a bike, typing on a keyboard, or driving a stick-shift.

These different memories communicate with each other, with the encoding of sensory and perceptual events making contact with LTM structures ("Was that Christopher?"), and the STM operating on LTM structures activated by current events and mentations ("What president followed Lincoln?"). People's ruminations can be sent back to LTM with embellishment ("I think she smiled at me!"), whether true or not. Some LTM memories flow into consciousness for personal examination (declarative memory), whereas other memories remain hidden below the level of consciousness yet remain capable of influencing a person's behavior (implicit memory).

The reason for all these memories is because research findings demand them. There is little correlation between STM and LTM, and thus, STM and LTM must be qualitatively different memory structures. How many digits or words a person can recite back in correct order before failing (STM) does not predict how many pictures a person can recognize 1 month after viewing (LTM). Variables that affect performance in an explicit memory task can be different from those that affect implicit memory.

Different brain mechanisms likely subserve these different memories. Damage to the hippocampus (located in the medial temporal lobe) destroys neither STM nor LTM. However, once damaged, no information entering STM can be passed onto permanent LTM storage, and episodic LTM stops at the time of hippocampal damage. Skills and habits are associated with the striatum; emotional responses and simple conditioning involve the amygdala. Damage from stroke can result in word aphasias (semantic memory deficits) that leave memories from the past (episodic) undisturbed, or the reverse. Positron-emission tomography (PET) scans reveal that different regions light up in response to semantic versus episodic retrieval. Therefore, the multimodal theory of human memory is embraced by most researchers.

There are other processes that operate in memory as well, called *control processes*. These include the operations that are applied to memories, like rehearsing a telephone number, forming a mental image, and attending to a message embedded within a complex of competing events. Some researchers have speculated that thinking and control processes are one and the same.

Episodic Long-Term Memories

Researchers often divide their investigation of episodic LTM into two broad categories: what makes people forget and what makes people remember. This simple division captures a number of topics of memory research, like state-dependent learning, eyewitness accuracy, autobiographical memory, flashbulb memories, imagination and memory, and so on.

Theories of Forgetting

Over the past 100 years, four major theories of forgetting have been proposed, broadly defined as *trace decay*, *trace interference*, *trace retrieval*, and *trace distortion*. Trace decay theory (formally called the theory of disuse) suggested that the passage of time itself was sufficient to cause forgetting. Literally, what is not used is lost, and the longer the passage of time was, the weaker the memory became. Although it is true that people's memory for more distant events will, almost without exception, be worse than their memory for more recent events, the passage of time

cannot be the sole (or even major) cause of forgetting. Rather, what happens during time is critical, and this discovery formed the basis for the next theory.

Interference theory, the dominant theory of forgetting from the 1930s to the 1970s, emphasized the importance of prior and intervening learning on forgetting. Two types of interference were proposed: retroactive and proactive interference. Retroactive interference is interference from learning that occurred between the original experience and the later recollection of it. For example, students who study American history in high school might receive a different interpretation of the same historical events from their college class. Some time after college, students may find that they can recall almost nothing of what was taught in their high-school history class—the intervening activity had "associatively unlearned" their memory of the original events.

Proactive interference is interference from learning that occurred before the event a person is trying to remember. The old adage "You can't teach an old dog new tricks" is a statement about proactive interference—the old habits keep interfering with the learning (or recollection) of a more recent habit (or event). If you cannot remember where you parked your car and find yourself at the location where you parked on a previous day, blame proactive interference. Interference theory proposed that all forgetting was due either to prior or intervening events. The tip-of-the-tongue phenomenon, the maddening experience of being unable to remember a particular fact you are certain you know, is thought to reflect the blocking of a memory by competing events. Later, once the competition has died down, the answer may flow back to you.

Encoding specificity theory came into prominence in the 1970s and continues today. Unlike interference theory, with its emphasis on prior and intervening events that block memory, this theory suggested that forgetting reflected the lack of a good retrieval cue. Find the correct cue, and a "forgotten" memory can be retrieved. Encoding specificity theory proposed that all experienced events are stored within a specific external and internal context. The external context included everything outside a person's body—the building the person is in, the sound of the overhead fans, the people next to him or her. The internal environment included the person's mood, thoughts, and mental state at that time. Because all events occur within a context, the context can become associated

with the event. Importantly, this context can provide the cues for successful retrieval of a forgotten event. If you witness a crime on the street, the cues from the physical environment and your concurrent mental state become (potentially) associated with your memory of the event. Returning to the scene of a crime (and perhaps having a beer, if you originally came from a bar) may help you remember details of the crime. State-dependent learning is part of the encoding specificity theory; it suggests that people learn in a particular state and that reinstatement of the original context, both physical and mental, can help retrieve a memory. Change the state, either physical or mental, and the memory may be lost; return to the state, and the memory flows into consciousness.

The memory literature is replete with marvelous examples of how contextual reinstatement and retrieval cues can be used to elicit an otherwise forgotten event. Usually, the more retrieval cues there are, the better the retrieval will be. Researchers in autobiographical memory have documented that, with enough good cues, people can retrieve events from many years past that would otherwise be lost. The importance of retrieval cues, lacking in interference theory, is the centerpiece of encoding specificity theory.

The fourth theory of forgetting, currently popular, is called overprinting, or memory distortion. Overprinting suggests that memories are not so much forgotten as they are distorted by (or blended with) subsequent activities. Elizabeth Loftus has demonstrated that a variety of postevent factors, such as hearing a later and different interpretation of a witnessed event, can distort, perhaps permanently, a person's original memory. Researchers have found other ways to distort a memory, such as repeatedly imagining the event in a different (and perhaps more pleasing) way. With enough imaginings, the original memory may be altered forever ("Perhaps I wasn't so cowardly after all!"). Another powerful way to distort a memory is to supply a false narrative of an event from your past, accompanied with an old photograph of you and others. With enough urging, you may eventually believe the narrative, even attaching a conspiratorial role to others in the photograph.

A remarkable demonstration of a false memory was reported by Nicholas Spanos. College undergraduates were contacted and told that they had been in an innovative experiment involving "visual enrichment" on later cognitive development. Following their birth, so the scam went, they were placed in a brightly

decorated room with multicolored mobiles that hung over their crib, and so forth. Some months later, and under the guise of a different experiment, these same students were asked to describe their earliest memories. Most of us cannot remember the initial 3 to 4 years of our lives (it's called "childhood amnesia"), yet 50% of these students recalled memories for events that occurred literally days after birth—bright yellow and blue colors floating overhead, and so on. These students maintained their hold on these earliest memories even when they were told that the visual enrichment experience was a bogus tale.

Separating reality from fantasy, once thought to be the hallmark of sanity, is likely confused to varying degrees in every person.

Factors That Enhance Long-Term Memory

What makes people remember is the flipside of forgetting. Some factors that enhance memory have been mentioned—reinstatement of context and providing retrieval cues that relate to a missing memory increases the likelihood that a memory can be retrieved. Other mundane practices also work, such as repetition and spacing, as in trying to remember the English meanings to Russian words. Levels of processing theory, proposed by F. I. T. Craik and R. S. Lockhart in 1972, suggest that deeper processing of a word, a face, or an event will lead to more durable memories than shallow (rote) processing.

Other powerful techniques involving mnemonic devices can also enhance memory. One technique uses natural language mediators. A natural language mediator is a translation, using language, that is applied to an otherwise low-meaningful event. For example, our license plate number, say NDZ-312, might forever escape our memory. However, translating NDZ-312 to something like "Naked Danish Zebras and there are 312 of them" might make it stick. Making up a poem or a sentence also works, using the first letter of each word as a cue. One example from music teachers is "Every good boy does fine." Each word begins with a letter that identifies one of the notes on the lines of the treble staff.

Weaving the events to be remembered into a simple story or narrative is effective. Gordon Bower demonstrated that subjects could remember 10 lists of 12 words, each studied for 1.5 minutes, at 94% accuracy if each list was first translated into a story. The control group, left to their own devices but with the same amount of study time, could remember only 14% of the words.

Many powerful memory aids involve mental imagery. The method of loci requires that you first identify a familiar route or set of locations (like a walk through campus from the psychology building to the student union), so that the sequence can be readily retrieved in order. The trick is now to form an interactive image for each successive word (or item) to be remembered with a location. At the time of recall, all you need do is to take a mental walk along the route and mentally inspect each location for the word residing there. The ancient Greeks used a much elaborated version of this technique so that speeches could be given without the aid of notes. Remarkably, the words used on one occasion can be mentally erased and the locations used again with a different list.

Other techniques use grouping or categorizing techniques. A classic study demonstrated that a college student could remember 79 consecutive digits by relating every three or four spoken digits to running events in track. This student took some 18 months of training to achieve this outcome, so it can take some practice, but it is a skill available to all. Presumably these techniques achieve their success by providing excellent retrieval cues at the time of recall, although why imagery works so well is less well understood.

Semantic Long-Term Memory

LTM also includes semantic LTM. This is the stuff a person knows but probably cannot remember when he or she learned it. All words are concepts, as are the natural categories (e.g., fruits, mammals) and higher-order concepts such as aesthetic style, disease categories in medical diagnostics, social stereotypes, and so on. Concepts are likely stored in semantic memory, so semantic memory effectively constitutes a person's memory of knowledge and facts. How people learn concepts and categories is one of the oldest issues in memory, dating back to the ancient Greeks, with Aristotle dividing the world into 10 mutually exclusive categories. Becoming an expert requires extensive training for particular concepts. Radiologists, who may have studied 100,000 x-rays, are experts in deciphering whether a particular x-ray likely contains one of myriad disease categories or nothing at all.

A fundamental principle of semantic memory is that concepts and relations are massively interconnected, so that activation of one concept or idea is

likely to prime other, related thoughts (activation theory). Theorists have developed complex representations of semantic concepts and their relations, such as multidimensional spaces. In these N-dimensional models, objects are represented as points in space, with related concepts nearby. Activate one concept, and the rings of excitation move outward to activate nearby ones, much like the expanding rings from a pebble dropped into a pond.

Future Directions of Research

Other lines of research and different disciplines involved in the study of LTM include the role of emotion on memory, how children learn and remember, and how memories from the different modalities are fused into a coherent memory. Memory scientists include cognitive psychologists, who study how much a person can remember and what makes a person forget; neurocognitive scientists, who study brain activity revealed by imaging techniques (PET scans and magnetic resonance imaging) in response to memory probes; and animal psychologists, whose findings with animals are often the first step to discovering important principles of memory.

What of Penfield and the perceptual moment hypothesis? Penfield's contention that people remember only a small fraction of what is stored is undoubtedly true. The jury is still out on his claim that people can uncover the "stream of consciousness" buried in memory, and he failed to distinguish real memories from memories fabricated and embellished by later mentations. Regardless, this century promises to reveal additional, extraordinary discoveries on the nature of memories.

Donald Homa

See also Episodic Memory; Memory; Short-Term Memory

Further Readings

Gabrieli, J. D. E., Desmond, J. E., Demb, J. B., Wagner, A. D., Stone, M. V., Vaidya, C. J., et al. (1996). Functional magnetic resonance imaging of semantic memory processes in the frontal lobes. *Psychological Science*, *7*, 278–283.

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind.* Chicago: University of Chicago Press.

Loftus, E. E. (2003). Make-believe memories. *American Psychologist*, *58*, 867–873.

Penfield, W. (1958). Some mechanisms of consciousness discovered during electrical stimulation of the brain. *Proceedings of the National Academy of Sciences*, *44*, 51–66.

Roediger, H. L., III, & McDermott, K. B. (2000). Tricks of memory. *Current Directions in Psychological Science*, *9*, 123–127.

Rubin, D. C. (1996). *Remembering our past: Studies in autobiographical memory.* Cambridge, UK: Cambridge University Press.

Schacter, D. L. (2001). *The seven sins of memory: How the mind forgets and remembers.* Boston: Houghton Mifflin.

M

The only man I know who behaves sensibly is my tailor; he takes my measurements anew each time he sees me. The rest go on with their old measurements and expect me to fit them.

—George Bernard Shaw

MAINSTREAMING

The term *mainstreaming* describes the inclusion of students with disabilities in schools and classrooms with nondisabled students. Since the mid-1990s, educators have tended to abandon that term for another, namely, *inclusive education.* The rationale for the change was to emphasize the active role that schools can play to effectively include students with disabilities in the academic and social life of schooling. Inclusive education refers less to a specific set of practices or place for schooling than to an orientation (i.e., a way of thinking), to an outlook or stance that emphasizes the full membership and active participation of all students, regardless of perceived ability/disability, race, class, or gender.

Although recognizing how individual differences affect student performance, an inclusive model suggests the importance of examining and refashioning broader social contexts of education. A paradox of the inclusive approach is that the very same factors that prove necessary for effective, quality inclusion are also those associated with quality education in general. Yet despite this seeming confluence of a specific policy agenda with a more universal goal, mainstreaming/inclusion has been unevenly implemented and is less available to certain segments of a school population.

Origins

The idea of mainstreaming in the United States grew out of the American experience with civil rights and the 1954 Supreme Court decision that separate schooling based on racial designation was inherently discriminatory; racial segregation led to students of color feeling less capable and less worthy than others and denied them access to full educational opportunity. The key principle borrowed from civil rights was the notion that children with disabilities are entitled to an education and that this should be with their nondisabled peers. The difference between mainstreaming policies that emerged and the classic *Brown v. Board of Education* decision is that the extent to which mainstreaming/inclusion would occur was left up to professionals to determine, possibly in negotiation with parents of children with disabilities.

In 1971, several families of children classified as mentally retarded filed a lawsuit in the Commonwealth of Pennsylvania, *PARC v. Commonwealth of Pennsylvania*, to establish the right of their children to receive an education. Actually, the case set forth a number of principles, included among them are the following five:

1. That the plaintiffs, all students classified as mentally retarded, had at great personal cost been denied access to public education

2. That all students could learn

3. That all students should be entitled to access public education

4. That education should be provided in the "least restrictive" setting possible; these parents asserted that students with disabilities were entitled to be included in classrooms with nondisabled students unless it could be shown to be impossible, even with specialized support services

5. That the education should be individually designed to enable students to succeed

Findings in the PARC Consent Decree set the stage for national policy fostering mainstreaming and, subsequently, inclusive schooling. In 1975, Congress passed Public Law (P.L.) 94–142, the Education for All Handicapped Children Act. This law was later amended and renamed the Individuals with Disabilities Education Act, P.L. 105–17. While P.L. 94–142 did not make reference to *mainstreaming* or *inclusion*, it did articulate a presumption that students with disabilities should be included with their nondisabled peers. Specifically, the law refers to the "least restrictive environment principle":

> To the maximum extent appropriate, children with disabilities, including children in public or private institutions or care facilities, are educated with children who are nondisabled; and special classes, separate schooling or other removal of children with disabilities from regular educational environment occurs only if the nature or severity of the disability is such that education in regular classes with the use of supplementary aids and services cannot be achieve satisfactorily. (20 U.S.C. Sec. 1412[5])

The term *least restrictive* derives from American legal reasoning that attempts to balance the exercise of state interests with protection of individual rights. Accordingly, the state should attempt to pursue its interests—for example, to provide an education to its children, including children with disabilities—in a manner that least intrudes upon or violates individual rights protected by the Constitution, such as liberty rights and freedom of association.

It is noteworthy that laws and regulations in Italy preceded the U.S. policy and were consistently less equivocal about inclusion. The Italians made inclusion a national standard, with no room for excluding any children. The Italian National Law 118, passed in 1971, mandates compulsory education for children with disabilities in regular classes of public schools, with no limitations or qualifications. Further law policies added that severity of disability should not prevent integration and outlined guidelines specifying class size, staff support, and number of children with disabilities per classroom. In comparison to the Italian policy, U.S. mainstreaming/inclusion policy leaves much room to argue about inclusion, particularly about how far it could go and for whom.

Mainstreaming in Practice

Early efforts to include students with disabilities in regular schools were not without difficulties. Mainstreaming was generally viewed as something that would occur student by student, individually and tentatively, without fundamentally altering the practice of maintaining a system of separate settings. The maintenance of special settings is actually required in law and is referred to as a *continuum*, ranging from the regular class to ever more segregated, disabled-only settings, from the resource room, to the special class, to the special school, and then to the residential institution or hospital.

In a national study of what was then called mainstreaming and later would be called inclusion, it was found that some programs actually perpetuated a certain amount of exclusion. Biklen and his colleagues described mainstreaming as taking these forms:

- *Teacher Deals.* This is where a school administrator agreed to allow individual special education teachers to place one or a few students out of the special class into the regular class. The drawback to this kind of arrangement is that it depends entirely on the goodwill of individual teachers, lacks institutional support and systematic planning, and is often abandoned when difficulties arise concerning how to adapt or modify the curriculum so that the student(s) can be successful.

- *Islands in the Mainstream.* This refers to a version of inclusion that is not very inclusive. A special class is located in a regular school, and the only interaction between students with and without disabilities is that which occurs in the lunchroom, during playground time, or in special subjects such as art and music. This approach is problematic precisely because it fails to achieve inclusion across the full school day and rarely leads to meaningful social or academic inclusion.

- *Unconditional, Purposeful Inclusion.* This refers to a form of inclusion where school administrators as well as teachers and parents plan together to make it work. The approach is pursued in such a way that difficulties are addressed as they occur. In this model, inclusion is viewed as a permanent characteristic of the school rather than as an experiment.

A decade later, another researcher found somewhat similar conditions in a more focused study. Examining inclusion involving children with Down syndrome, Kliewer presented three different status positions that students with Down syndrome occupy as analogous to foreigners in a new land:

1. Alien—the student is always treated as an outsider and is basically kept separate or excluded

2. Squatter—the student occupies a space, but is always subject to being removed or displaced

3. Citizen—here the student enjoys the right of belonging and of participating in all aspects of classroom and school life

In this model, the idea of achieving inclusion is adopted as a central concern of the school, much as one might embrace the goal of teaching mathematics or literacy. Instead of seeing inclusion as something to be achieved only by those students who prove themselves worthy of it, purposeful inclusion/citizenship imagines the school engaging in reforms that create a context where a student is fully immersed in the life of the school.

Questions About Inclusion

Many questions arose as teachers and faculty considered implementing inclusion. The dominant one was this: Is inclusion a good idea? Bogdan and Kugelmass answer this question by challenging the question itself, noting that asking whether inclusion is a good idea is a bit like asking, Is Tuesday a good idea? Their point is that the term *mainstreaming* exists, but what it actually looks like and the results of pursuing it will be determined not in some abstract, objective sense, but by engaging in inclusive practice, studying it, and adapting strategies in the process.

Among the many questions that have been raised about inclusive schooling are the following: Will it be more expensive than segregated services? Will it require specialized training for which most teachers are ill prepared? Is there research to demonstrate that

it works for students with certain disabilities but not for others? Of course, this last question could be framed the other way around, such that if one wanted to forcibly exclude children on the basis of disability research, one would have to prove that the educational results of segregation were substantially superior; otherwise, how could a limitation on freedom of association be justified? Early research on each of the above questions appeared to support mainstreaming. Mainstreaming/inclusion appeared to minimize the difficulty that students with severe disabilities might experience in generalizing skills learned in one setting to another; students with mild disabilities could be served equally well in regular classes with support as in resource rooms, and students' acceptance of each other appears to grow through proximity.

More recent analyses of inclusive education are similarly supportive of the inclusion agenda, even at the secondary level, where it is generally thought to be more difficult and less widespread. Studies examining the attitudes of general education teachers toward the inclusion of students with disabilities suggest that support for the practice of inclusion improves with training in specific methods related to it. And studies that examine the relative benefits of inclusion over segregation for matched groups tend to demonstrate improved results for the included students. For example, Fisher and Meyer compared 40 students in two groups, self-contained versus inclusive, and found that the inclusive group made greater gains after 2 years on both academic and social scales. Using a posttest only, control group design, Kennedy, Shikla, and Fryxell compared two matched groups of middle school students on measures of social interaction, social support behavior, and friendship networks. The students in the inclusive setting reported significantly higher levels of contact with peers than students in the self-contained setting. In addition, they had larger friendship networks and received and provided higher levels of social support.

Ryndak, Morrison, and Sommerstein explored engagement with literacy activities for a young woman who began her academic career in self-contained classes and moved to an inclusive setting after 10 years. This 7-year case study provides compelling evidence of growth in language and written literacy after moving to a classroom with typical peers. Using a similar comparison-across-settings approach, Dore, Dion, Wagner, and Brunet observed two 15-year-old students in a self-contained classroom from September to March and then after their transfer to a regular classroom for the

remainder of the school year. In the inclusive settings, the students were less likely to be engaged in lecture and activities similar to their peers and more likely to be working on their own individual activities. However, few modifications were made other than the presence of a teaching assistant and limited tutoring. In a time sample study of students with mental retardation labels who spent at least part of their day in regular education classrooms, Agran found that students with disabilities were more likely to be working on standards-based tasks in inclusive settings when compared to segregated settings, where they were more likely to be working on below-grade-level standards or Individualized Education Program goals. Taken together, these studies demonstrate possibly mixed results, which would confirm the argument that the nature of programming associated with mainstreaming/inclusion matters and that the meaning of mainstreaming/inclusion cannot be understood apart from evidence about the specific practices done in its name.

Race, Class, and Inclusion/Exclusion

Ironically, in the years that have followed the 1954 *Brown* desegregation decision, special education has been used to reinstate patterns of segregation. In Washington, D.C., as revealed in *Hobson v. Hansen* (1967), tracking of students followed racial lines such that White students were typically placed in high-track, or advanced, classes and Black students were consigned to lower tracks and to special education. Although the *Hobson*-era situation predated policies that give preference to inclusion, the patterns of segregation via special education placement revealed in *Hobson* persist and are reflected in national data on inclusion. The 26th Annual Report to Congress Office of Special Education Programs (2004) finds that

- Black students are 3.04 times more likely to be identified as mentally retarded and 2.25 times more likely to be identified as emotionally disturbed than all other ethnic groups.
- 28.5% of classified Black students spend more than 60% of the school day segregated in disabled-only programs in comparison to 14.7% of classified White students.

Ferri and Connor note that Black males are two times more likely to be labeled mentally retarded in 38 states. Students of color are several times more likely to be labeled either mentally retarded or

emotionally disturbed than are students from all other ethnic backgrounds. Clearly, there is little consistency across state lines in terms of segregation and integration practices. The general conclusion one can draw from the national data on race and placement is that the inclusion movement bypasses students of color. Put another way, historic patterns of racial discrimination, including segregated schooling and inadequate teaching, persist and may be aided by special education as currently fashioned.

Principles of Inclusion

Much of the mainstreaming/inclusion literature focuses on approaches or strategies. Below are some of the concepts that have emerged in the literature. Essentially, these are the ideas of people who have been active in conducting research on and formulating demonstrations of mainstreaming/inclusion practice; they are concepts that are being used to frame mainstreaming/inclusion work. Although not an exhaustive list, it provides a sense of current goals and practices of inclusion.

Presumption of Competence

This assumes that all children are educable and that the best approach to educating any child is not to presume incompetence. When one presumes competence, the onus of responsibility for making education work shifts to the teacher and to school administrators, parents, and others. If a child appears not to be learning, but is nevertheless presumed competent, then educators must ask themselves what they can do differently in order to see the competence. This is similar to the concept of the Least Dangerous Assumption introduced by Donnellan. Even in the absence of evidence, educators should adopt practices and policies that are least likely to limit potential. According to this principle, for example, it is dangerous to assume that a student must be segregated, for this may limit the child from having access to the typical academic curriculum and to interaction with typical peers. The less dangerous assumption is that a child may benefit from inclusion.

Purposeful Inclusion and Citizenship

Ideally, educational leaders join with teachers and parents to make inclusion work. And ideally, it will be adopted not as an experiment—the danger of treating it

as an experiment is that when difficulties arise, there may be a tendency to abandon the experiment rather than to do the hard work of trying other strategies—but as a favored, even essential approach. In such a model, inclusion will be regarded as purposeful, where the student with a disability is seen as a full citizen in a classroom of peers, and not as an outsider who is still waiting for his or her immigration status to be transformed from alien or squatter to citizen. The outstanding examples of effective inclusion have occurred in schools that have identified inclusion and/or belonging as central goals, alongside the commitment to educate both boys and girls or to honor equality for ethnic and cultural minorities. In addition, schools that embrace mainstreaming/inclusion often make a point of not labeling students in public, everyday discourse. Students are known by their names, their personalities, their interests, and so on, rather than by the kinds of support they require or by diagnostic labels.

Natural Proportions

To avoid overrepresentation of students with disabilities, Brown suggests that a useful principle is to create inclusive classrooms where disability is present in "natural proportions"; in other words, to the extent that would be expected in a normal distribution within society at large. Thus, children with disabilities would expect to attend the schools they would have attended if not disabled.

Development of Local Understanding of Students

A key skill for teachers is observation. Sometimes referred to as qualitative research, unobtrusive research, phenomenology, grounded theory, inductive inquiry, or participant observation, this research method examines how children interact with and interpret the worlds they encounter, and therefore how they learn. It is also helpful to ask parents and teachers to describe those situations where the student appeared to demonstrate complex understanding of something. Invariably, they are able to point to numerous examples, such as when a child with cerebral palsy laughed before other students at a complex joke. For educators, recognition of such instances where students demonstrate complex understanding is important, for they can extrapolate from these and attempt to replicate those conditions throughout the curriculum.

Universal Design

A concept borrowed from architecture, Universal Design is the idea that instead of adapting a curriculum that was designed for a narrow, "normal" segment of the population, the best way to design instruction and curriculum is to consider the needs of the broadest spectrum of the population at the outset, such as a student who has differences in expressive language or in receptive language, or differences in mobility, hearing, or sight. Through universal design, supports are available to a student in a way that appears normative or natural, even elegant.

Multiple Intelligences

It is widely recognized that students have different ways of learning. Some do well with lectures, whereas others seem to make sense of lecture content only if it is accompanied by hands-on activities or projects. Others need visual displays of information along with spoken representations. Gardner describes these diverse ways of learning as multiple intelligences and recommends that teachers consider how to deliver the curriculum using all of the intelligences.

Building on the Literacies That Students Bring to the Classroom

All students come to school having already learned a great deal from family members and the community. Many of the classic examples of effective teaching, such as that described by Ashton-Warner about her work with Maori children in New Zealand or by Ladson-Billings in her book *The Dreamkeepers*, place emphasis on how teachers can build on the knowledge that students bring with them when they enter the classroom. The role of education, then, is to create curricula that take advantage of prior experience or allow it to be revealed, and to foster productive inquiry and reflective discourse.

Authentic Assessment

Routinely, schools and education systems evaluate students by using standardized assessments. These may or may not measure student performance in relation to what they have been exposed to in class and what educators wanted them to learn. Authentic or criterion-referenced assessment refers simply to evaluations that

use the materials of the day-to-day curriculum to assess student learning.

Listening to Affected People

Few resources have been as valuable to educators as the firsthand narratives of people with disabilities. Through these accounts, educators learn about how students have interpreted the educational strategies they experienced as well as their own self-invented strategies. This principle recognizes that treating all students the same could be discriminatory toward some if it means that they cannot access the curriculum.

Cooperative Instruction

Even though students demonstrate dramatically different skills and knowledge, they often benefit from collaboration. One way to achieve this is through what has been called *cooperative learning* or *cooperative instruction.* The key principles to make this work are to ensure that all students have an individual role, that the total project result depends on each student's contribution, and where the individual student contribution builds on and expands on a student's abilities and knowledge.

Access to Academic Curricula With Appropriate Supports

It has sometimes surprised educators to learn that even when students with certain disabilities were unable to demonstrate their comprehension of the curriculum, often the benefits of being in the classroom where the traditional academic content was being taught became apparent several years later. That is, students were later able to show that they acquired knowledge and skills incidentally. However, simply granting access isn't enough—educators have to provide the supports needed for students to participate and show what they know.

Friendships (Circles of Friends/Alliances)

It is well known that students' tolerance and acceptance for each other and each other's differences grows when they are in proximity with each other and engaged in shared activities. Recent research demonstrates that with structured opportunities for interaction and modeling by teachers and others, students with and without disabilities learn to interact with each other and can develop close personal friendships.

Scaffolding in the Zone of Proximal Development

Educators' applications of the Vygotskyan concept of zones of proximal development involves examining a child's intellectual participation and hypothesizing next steps that will stretch that engagement from the skills a student demonstrates currently to imagined next steps, all with teacher, peer, and material/curriculum mediation. In addition, educators need to be always thinking about the next environment for the students. One consideration is what academic and social skills will be expected of students when they enter the next level of education.

Douglas Biklen and Christine Ashby

See also Cultural Diversity; Inclusion; Least Restrictive Placement; Special Education; Students' Rights

Further Readings

Ferri, B., & Conner, D. (2006). *Reading resistance.* New York: Lang.

Jackson, L. (2002). *Freaks, geeks and Asperger syndrome.* New York: Jessica Kingsley.

Janney, R., & Snell, M. E. (2006). *Social relationships and peer support* (2nd ed.). Baltimore: Brookes.

Jorgensen, C. (1998). *Restructuring high schools for all students.* Baltimore: Brookes.

Kliewer, C., Fitzgerald, L. M., Meyer-Mork, J., Hartman, P., English-Sand, P., & Raschke, D. (2004). Citizenship for all in the literate community: An ethnography of young children with significant disabilities in inclusive early childhood settings. *Harvard Educational Review, 74,* 373–403.

Kluth, P., Straut, D. M., & Biklen, D. (2003). *Access to academics.* Mahwah, NJ: Lawrence Erlbaum.

Oyler, C. (2006). *Learning to teach inclusively: Student teachers' classroom inquiries.* New York: Teachers College Press.

Sapon-Shevin, M. (1999). *Because we can change the world: A practical guide to building cooperative, inclusive classroom communities.* Boston: Allyn and Bacon.

Shapiro, A. (2000). *Everybody belongs.* London: Routledge.

Tomlinson, S. (2004). Race and special education. In L. Ware (Ed.), *Ideology and the politics of (in)exclusion* (pp. 76–88). New York: Lang.

MALNUTRITION AND DEVELOPMENT

Unless a more specific usage is indicated, *malnutrition* generally refers to *protein-energy malnutrition (PEM).* The term was introduced because the two

deficits tend to occur together; an isolated protein deficiency is rare. Children are particularly vulnerable to malnutrition, and severe malnutrition is associated with high mortality. But many children live through chronic mild to moderate malnutrition, and there is reasonably good evidence that it adversely affects their cognitive development. Micronutrient deficiencies are also widespread. Two micronutrient deficiencies that also affect cognitive development are iodine and iron deficiency.

As well as the type of malnutrition, its timing, duration, and severity are important. Malnutrition may be prenatal or postnatal, and after birth it may be found in early or late infancy, childhood, or adolescence. Its origins can be equally varied. They can include contributions from the availability of food and its composition, the care provided by parents and others, the child's appetite and food preferences, and the child's illnesses and disabilities. The ability to eat is often taken for granted, like the ability to talk, but like talking, eating also involves very complex motor skills, and where they are compromised, as they are, for example, in children with cerebral palsy, serious malnutrition can result.

For obvious reasons, children have not been deliberately malnourished, so research in this area has principally used observational studies. But the families of malnourished children tend to differ from the families of well-fed children in many other ways that also affect their development, so the inferences that can be drawn from observational studies depend on the extent to which these can also be measured and taken into account. A valuable additional source of information comes from studies in which additional nutrients have been provided for children in trials that allow a later comparison with unsupplemented control children.

Protein-Energy Malnutrition

Protein-energy malnutrition is usually identified by its effects on growth, and the most commonly used measures of growth are height, weight, and weight relative to height. These measures can be interpreted only in relation to norms or standards, which are summaries of the distribution of the measured values in populations, often in the form of a growth chart. In general, the growth of *well-nourished* children in different populations is sufficiently similar for it to be possible to use a single set of standards internationally, as recommended by the World Health Organization (WHO). The United States National Center for Health Statistics/WHO international

reference standard has mostly been used for this purpose. Criteria for malnutrition can be expressed in centiles (e.g., a height below the third centile of the reference population) or in standard deviation scores (e.g., a height more than two standard deviations below the average).

Using this criterion, about a third of the children in developing countries were malnourished in 2000 (183 million). This proportion declined between 1980 and 2000 in most areas of the world (but not in Eastern Africa). There is good evidence that protein-energy malnutrition in the childhood years is associated with slower development, as measured using well-standardized scales such as the Bayley Scales. There is also evidence that it is associated with lower intellectual ability at school age. This association has been found with intelligence tests and with more specific tests of psychological functions, such as attention and working memory, and of educational attainment, for example, in reading. The mechanism by which malnutrition has these effects is still uncertain. It could involve structural changes in the brain, or changes in the behavior of the child leading to reduced interaction with the environment, which might additionally result in lower levels of stimulation from adults or other children caring for them (this is the *functional isolation* hypothesis).

It is difficult to make useful generalizations concerning the size of the effect of protein-energy malnutrition on intellectual development internationally. It tends to be associated with other kinds of malnutrition and with other adverse circumstances. There is reasonably good evidence that the size of the effect depends on the timing of the malnutrition, with poor nutrition in the earlier stages of life having a bigger effect, and on the duration of follow-up, with effects that are present in early childhood tending to "wash out" later (although there is some evidence that they may become more important again when general cognitive abilities decline in old age).

In industrialized countries, most of the available work on cognitive outcomes of malnutrition has been carried out in relation to children who have "failed to thrive." This term is used for infants or young children whose weight gain is poor. Criteria similar to those used to identify malnutrition (say, a weight gain in the lowest 5%) are used to identify it. Its immediate cause is likely to be poor nutrition, perhaps with some contribution from ill health, and it is therefore likely to have effects similar to those of malnutrition in other parts of the world. Well-documented sequelae include

developmental delay in infancy and some reduction in intellectual ability at school age (about three IQ points). There is less direct evidence at the moment that school performance is affected, although that may simply reflect a lack of sufficiently large studies.

Iodine Deficiency

Iodine deficiency is traditionally found in populations in which iodine levels in the soil are low, leading to low iodine levels in food crops. Its prevalence is reduced to some extent by the import of foodstuffs from iodine-rich areas, and to a much greater extent by the iodization of salt, which essentially eliminates the problem, as it has in the Americas. Nevertheless, recent estimates suggest that 285 million children worldwide (36.5%) currently have an inadequate iodine intake. Iodine is needed for the synthesis of thyroid hormones, which have an important role in normal brain development, so iodine deficiency before and immediately after birth has particularly adverse effects. The syndrome that was traditionally referred to as "endemic cretinism" is a consequence of iodine deficiency in the fetus and is associated with major learning disabilities, but lesser degrees of iodine deficiency before birth also impair fine motor skills and intellectual abilities, and later iodine deficiency is also important. In children of school age, it is associated with lower intellectual abilities as assessed, for example, with Raven's Progressive Matrices, and with poorer school performance in areas such as spelling and reading. The best estimate currently available suggests that IQ is reduced by an average of 13–14 points in communities that are chronically iodine deficient.

Iron Deficiency

Iron deficiency can take the form of iron deficiency anemia, or of milder states of iron deficiency that do not affect hemoglobin production (iron deficiency without anemia). The most common causes of iron deficiency in infancy and early childhood are a poor dietary intake of iron, and loss of blood from the gastrointestinal tract as a result of gut parasites or cow's milk intolerance. An influential review in the 1980s suggested that about half the world's children under 5 years of age were anemic. There is considerable evidence for a progressive reduction in iron deficiency and anemia in young children in the United States from the early 1970s as a result of increased

breastfeeding, delaying the introduction of cow's milk, the fortification of infant formulas and cereals with iron, and the Special Supplemental Food Program for Women, Infants and Children (WIC). But even in the United States, recent figures suggest a prevalence of iron deficiency in 9% of children under 2 years of age and 2% to 3% in children from ages 3 to 11. About a third of these children are anemic.

Iron deficiency anemia is consistently associated with developmental delay in infancy and early childhood. The association has been found in studies in which reasonable efforts have been made to control for other relevant factors, including studies in industrialized countries. Children who are anemic as infants tend to do less well on cognitive tests and on measures of school achievement. There is reasonably good evidence that in older children, these adverse effects can be ameliorated by iron treatment. Its effectiveness in younger children is less clear.

The Adequate and the Best

Following a traditional way of thinking, the effects of malnutrition dealt with so far are associated with the classically identified nutritional deficiency disorders. But simply avoiding deficiency disorders may not be enough to guarantee optimal nutrition for a child.

A low birthweight, for example, is traditionally defined as a birthweight below 2,500 grams and can be due to a delivery before term or to poor growth before birth. Although low birthweight at term has a number of different causes, prenatal malnutrition is certainly one of them, and maternal malnutrition is the most important cause of low birthweight in infants born at term in developing countries. Many studies have examined the intellectual effects of being born below 2,500 grams, and IQ is somewhat lower than average in these low birthweight children. But this is only one manifestation of a more general relationship between a higher birthweight and a higher later IQ, which increases systematically with birthweight up to about 4,000 grams. The overall increase is about 5–10 IQ points. If this is a nutritional effect, it suggests that optimal prenatal nutrition, at least as regards intellectual potential, is more than is needed simply to avoid a low birthweight as traditionally defined.

Another example involves the fatty acids arachidonic acid (AHA) and docosahexaenoic (DHA) acid. These are synthesized in adults, but infants rely more

on AHA and DHA from their diets. Fatty acids have specific structural roles in cell membranes in the brain and the retina. They come from the mother before birth. They are also present in breast milk but are not present in substantial amounts in formula milk as it was traditionally manufactured, and there is direct evidence that breastfed infants have had significantly greater concentrations of DHA in brain tissues. These facts have led to extensive studies of the effect of supplementing bottle-fed infants, especially those born preterm, with AHA and DHA. There is evidence that such supplementation improves visual acuity in the short term, and some studies suggest that it benefits cognitive development as well, although the results are not entirely consistent.

Implications

There are many good reasons for ensuring that children are well fed, and the adverse effects of poor nutrition on intellectual development are among them. Except for prenatal effects of iodine deficiency, these adverse effects are not very large, but they are likely to be very widespread and to be associated with other risk factors that also have adverse effects. The effects of malnutrition on cognitive development in children appear to be rather general and may operate by multiple pathways. Although progress is being made with the neuropsychology of this area, the evidence across studies that malnutrition affects specific psychological functions is not sufficiently clear to be practically helpful. As far as we know at the moment, all those features of an educational system that improve the attainments of other children are also likely to improve them in children whose early nutrition has been poor.

R. F. Drewett

See also Intelligence and Intellectual Development; Mental Retardation; Percentile Rank; Poverty; Risk Factors and Development

Further Readings

Drewett, R. F. (2007). *The nutritional psychology of childhood.* Cambridge, UK: Cambridge University Press.
Grantham-McGregor, S., & Ani, C. (2001). A review of studies on the effect of iron deficiency on cognitive development in children. *Journal of Nutrition, 131,* 649S-668S.
Schmidt, A. T., & Georgieff, M. K. (2006). Early nutritional deficiencies in brain development: Implications for psychopathology. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology* (Vol. 2, pp. 259–291). Hoboken, NJ: Wiley.
Simeon, D. T., & Grantham-McGregor, S. M. (1990). Nutritional deficiencies and children's behaviour and mental development. *Nutrition Research Reviews, 3,* 1–24.

Maslow's Hierarchy of Basic Needs

Abraham Harold Maslow's theory of human motivation postulates that all human beings, regardless of culture, have basic needs that can be arranged on a hierarchy according to prepotency or pressing drive for gratification. His earlier writings presented five basic sets of needs (or need groups), and he later added two additional need groups. From the lowest level of needs (the most prepotent needs) to the highest level, these include physiological needs, safety needs, belongingness and love needs, esteem needs, cognitive needs, aesthetic needs, and need for self-actualization (see Figure 1). Maslow classified the four lowest need groups on his hierarchy as *deficiency needs* (physiological, safety, belongingness and love, and esteem) and the three highest need groups as *growth needs* (cognitive, aesthetic, and self-actualization). Table 1 presents a summary of the hierarchy of needs by name and description of need group, ordinal level on the hierarchy, and category (deficiency needs vs. growth needs).

Deficiency Needs

The most prepotent need group, physiological needs, relates to the body's need for food (hunger), water (thirst), air (oxygen), sleep (rest), and optimal temperature (comfort) in order to survive and maintain a state of physiological homeostasis or equilibrium. Safety needs, the second most prepotent need group, include needs for security; protection; stability; and freedom from harm, fear, or constant anxiety. Belongingness and love needs, the next level, are described as the need to belong to and feel loved by significant others and one's affiliated groups, such as family, neighbors, friends, fellow job employees, social club or fraternity members, gang members, or, in general, one's own primary group. The esteem needs, the next hierarchical
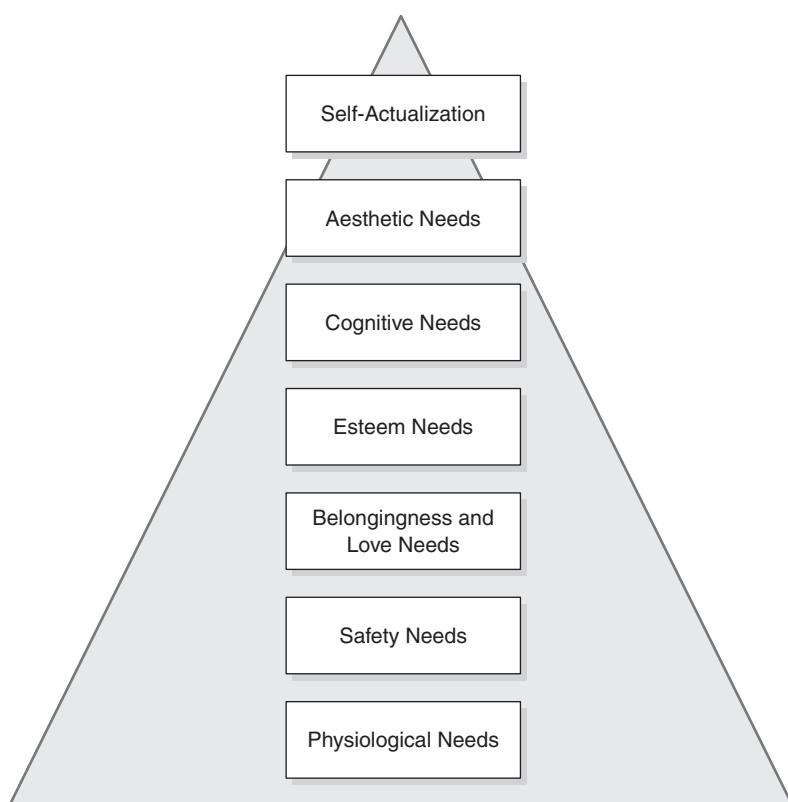
Figure 1   Abraham Maslow's Hierarchy of Basic Needs

level, have to do with self-esteem and deserved esteem from others, that is, based on one's accomplishments, status, or appearance. Moreover, the esteem needs encompass the need for approval, self-respect, and respect from others.

Growth Needs

Cognitive needs, the first growth need group on Maslow's hierarchy, are described as needs to know, understand, and explore one's environment and world. Aesthetic needs, the second growth need group on the hierarchy, are defined as needs to appreciate, seek, and strive for beauty, symmetry, and order in the world and in one's environment. The highest Maslowian need is self-actualization, which is the need to develop one's common potential and unique talent at the highest possible level of growth and achievement. It is a personal need for growth and fulfillment toward becoming all that one can become—toward becoming a healthy person.

The Nature and Dynamics of Maslow's Basic Needs

Maslow theorized that needs explain much (but not all) of human motivation and striving, and that the gratification of basic needs leads to a holistically healthy or self-actualizing personality, whereas a lack of gratification of the basic needs results in psychopathology or sickness. Maslow used the term *sick* to include biological, psychological, and spiritual illness, or holistic illness vis-à-vis holistic health.

The most prepotent needs of the person occupy a conscious effort and striving for gratification or satisfaction, whereas the less prepotent needs are unconscious, minimized, denied, or suppressed. Therefore, when one need group is satisfied, the next prepotent need group emerges to dominate the drive or conscious motivational efforts of the person. Moreover, the gratification of the four deficiency needs (physiological, safety, belongingness and love, and esteem) are necessary before the human organism can sufficiently focus on self-actualization and other growth needs as a means toward a healthy or superior personality. Gratification of a need does not imply total gratification or 100% satisfaction, but it suggests the sufficient gratification of one need group before the person focuses on the next need level.

Maslow believed that the basic needs are "instinctoid," that human beings have a natural, innate tendency toward gratifying basic human needs, growing, and expressing goodness versus evil. Just as positive societal conditions can facilitate need gratification, negative societal conditions and circumstances can thwart the healthy gratification of needs and create a repressive, fearful, and self-diminishing tendency for some human beings. In addition, very few human beings become self-actualizers, and this level is most likely to occur for adults and older persons and least likely or unlikely to occur among children.

Metaneeds

Maslow states that metaneeds are synonymous with metavalues or B-values (being values). The B-values behave like needs; therefore, Maslow also referred to

Table 1   Basic Needs by Category, Level, and Description

Category of Needs	Basic Needs (Lowest to Highest)	Description of Need Group
Growth Needs (Being Needs or B-Needs)	Self-Actualization	Need for growth to develop one's common and unique potential or talent; to find one's mission, purpose, or vocation in life; need for fulfillment
	Aesthetic Needs	Need for beauty, order, and symmetry
	Cognitive Needs	Need to know, understand, and explore one's world
Deficiency Needs (D-Needs)	Esteem Needs	Need to gain respect and recognition from others; need for self-esteem and prestige for one's accomplishments and appearance; need to feel a sense of adequacy and self-enhancement
	Belongingness and Love Needs	Need for acceptance and approval of others; need to belong to a group and acquire warmth from another or others; need to love and be loved
	Safety Needs	Need for a sense of security and freedom from threat or danger; need for protection from harm (both physical and emotional), need for adequate shelter for protection against the atmospheric elements
	Need for Physiological Maintenance	Physiological needs for survival and bodily maintenance (e.g., food, water, air, sleep, and optimal environmental temperature)

Note: Maslow referred to the basic needs or need groups as *sets of needs*, suggesting that a need group often contained subneeds or multiple needs.

them as metaneeds. Metaneeds or B-values are prevalent among self-actualizing people and those who choose a path of growth. They include the need or value for

- *truth* (honesty, reality, purity, completeness);
- *goodness* (rightness, justice, benevolence);
- *beauty* (form, perfection, uniqueness, honesty);
- *wholeness* (unity, integration, oneness);
- *dichotomy-transcendence* (acceptance, resolution, transcendence of opposites);
- *aliveness* (spontaneity, full-functioning, opposite of deadness);
- *uniqueness* (individuality, noncomparability, novelty);
- *perfection* (nothing lacking, everything in the right place, completeness);
- *necessity* (inevitability, it must be that way, it is good that it is that way);
- *completion* (finality, closure, fulfillment of destiny);
- *justice* (fairness, oughtness, nonpartiality);
- *simplicity* (honesty, unmistakability, nothing extra or superfluous);
- *richness* (totality, nothing missing or hidden);

- *effortlessness* (lack of strain, grace, perfect and beautiful functioning);
- *playfulness* (fun, joy, humor); and
- *self-sufficiency* (autonomy, independence, self-determination).

Implicit in the metaneeds or B-values, as well as the basic human needs, is Maslow's presupposition that the core of human nature is good rather than evil and growth-oriented rather than deficiency-oriented. Moreover, a lack of ability to fulfill metaneeds can result in *metapathologies* or minor psychological disorders.

Society, Culture, and the Dynamics of Basic Needs

During early publications of Maslow's hierarchy of basic needs, there was subsequent criticism regarding the rigidity of the fixed hierarchical levels and the theoretical prerequisite to fulfill needs in a fixed order, that is, one need group after another. Many of the questions were based on the fact that although the

basic needs were seemingly for all human beings, people differed in how they fulfilled such needs due to individual differences and cultural differences or cultural worldview. Therefore, in his later writings, Maslow revised the theory of basic needs and acknowledged the need for cross-cultural research in his theory of motivation. He addressed instances in which some persons, because of individual differences, may not follow the theoretical order of gratification of the basic needs; for example, such persons may give preferential attention to a higher need before attention to or sufficient gratification of the preceding, lower need on the hierarchy. For example, self-actualizers may almost totally ignore lower needs in order to practice and achieve perfection in performance or to create a scientific discovery or new product. As an example, famous musical composer George F. Handel refused food from his servants and had little to no human contact while locking himself in his room to write *The Messiah* (a massive 252-page book of music) in only 24 days. Along the same line of deferring lower needs, Maslow explains that some persons ignore love for the fulfillment of the higher need of esteem; therefore, greater potency emerges for esteem before adequate attention is given to the gratification of the lower need for love and belongingness.

Recognizing societal variations across cultures and countries, Maslow discussed preconditions of society that could either facilitate or limit an individual's ability to satisfy the basic needs. Some of the preconditions that can facilitate need gratification include (a) freedom of speech, (b) freedom of choice, and (c) freedom from injustice. As can be expected, need gratification and self-actualizing possibility are limited in countries or cultures that oppress or violate the human rights of various groups or individuals. Throughout the world and across historical eras, there have been particular groups and individuals whose need satisfaction and growth toward self-actualization have been blocked or stifled due to injustice, oppression, and exclusion from opportunity and resources.

Raised and educated in U.S. culture, Maslow was a product of Western psychological training; therefore, his theory represents an emphasis on individual needs and motivation (a focus of Western cultures) versus group needs or group-related motivation (e.g., a value of Eastern or Asian cultures). Although Maslow's basic needs seem to be universal to all human beings, the question is whether there are differences in the hierarchical order of need emergence,

manner and level of gratification, and the importance of values attached to particular needs as related to the person's cultural worldview and context.

Self-Actualization

Maslow acknowledged and credited Kurt Goldstein for coining the term *self-actualization;* nonetheless, it was Maslow who popularized and gave broader meaning to the concept. He viewed self-actualization as a growth need and a process toward becoming a psychologically healthy personality or superior human being in terms of organismic functioning at the highest level of personal fulfillment, human functioning, and enhancement of the self. In other words, Maslow believed that an artist must paint, a gifted musician must create and play music, and a poet must write; in other words, what a person is gifted to do, that person must do in order to be happy, fulfilled, and actualized.

Although Maslow viewed self-actualizing people as being creative in thought and productivity, he eventually concluded that highly creative persons and self-actualizing or healthy persons were not necessarily the same; rather, some persons who performed at a high level of creativity were not necessarily mentally healthy or healthy personalities, whereas others were mentally healthy or self-actualizing people. Nevertheless, the gratification of higher human needs or growth needs was theorized by Maslow as the natural passage to a healthy personality or to positive mental health, and failure to gratify basic needs in a sufficient manner could result in pathology or mental illness.

Characteristics of the Self-Actualizing Person

Self-actualization is an ongoing process and not an end within itself. Although no one is expected to achieve the perfect state of self-actualization or gratify this need 100%, there are a few human beings who rise to the hierarchical need level of self-actualizers or self-actualizing people, that is, by developing their potential at a high level and exhibiting most of the self-actualizing characteristics that Maslow identified. These characteristics of self-actualizing people are based on Maslow's study of historical figures and personal acquaintances. A summary of the 15 characteristics of self-actualization follows:

1. *Superior Perception of Reality.* Perceives and judges situations and people efficiently and accurately; sees the world as real and from a logical perspective; is unthreatened by difference or by the unknown.

2. *Acceptance of Self, Others, and Nature.* Accepts oneself without unrealistic shame, guilt, complaint, or anger; accepts others and people as they really are; accepts self as a healthy individual (e.g., accepts one's sexuality and shortcomings).

3. *Spontaneity.* Is spontaneous and unconventional in behavior as influenced by inner thoughts and personal ethics.

4. *Problem Centeredness.* Focuses on problems outside of oneself instead of oneself as a problem; is occupied with purpose, mission, and a problem to be resolved.

5. *Detachment.* Can be solitary and, at times, prefers solitude and privacy more than the average person; is not dependent on others but rather depends on self; is not exceptionally disturbed by misfortune or setbacks.

6. *Consistent Appreciation.* Can appreciate the same experiences and encounters with equal or greater enjoyment and intensity.

7. *Mystic Experiences/Oceanic Feeling.* Is capable of reaching frequent peak experiences or ecstatic, spiritual experiences of unlimited heights—peak experiences that are characterized by a feeling of transforming strength and a precious sense of well being and purpose in life.

8. *Identification With Mankind.* Identifies with humankind and tends to feel deep compassion for the suffering of human beings; is interested in improving the human condition.

9. *Interpersonal Relations.* Shares few close or intimate relationships; however, handles many superficial relationships effectively.

10. *Democratic Character.* Establishes a natural relatedness to various ethnic, cultural, religious, and educational groups; is tolerant of group and individual differences; and is open to learning from anyone, but is selective in choosing relationships with others.

11. *Discrimination Between Means and Ends.* Focuses on ends to the subordination of means, especially ethical ends as well as ends or goals related to personal growth and mission.

12. *Independence of Environment.* Is less dependent on other people or culture, or extrinsic satisfaction and more dependent on continued growth and development of inner potential.

13. *Philosophical Sense of Humor.* Possesses a philosophical and spontaneous sense of humor that is not degrading of or hostile to others; does not laugh at the miseries or weaknesses of others.

14. *Creativeness.* Is creative, original, and divergent in thoughts and actions.

15. *Resistance to Cultural Conformity.* Resists blind conformity to culture, especially conformity that interferes with inner ethical principles or personal growth and fulfillment.

Studies of Self-Actualizing People

Maslow studied a small sample of personal acquaintances and friends as well as historical figures. Among the historical figures whom he found to exhibit characteristics of self-actualizers were Ludwig von Beethoven, Albert Einstein, Benjamin Franklin, Sigmund Freud, Mahatma Ghandi, William James, Thomas Jefferson, Abraham Lincoln, Eleanor Roosevelt, and Henry David Thoreau. Also, Frederick Harper studied three internationally known African American protest leaders and found Frederick Douglass, Martin Luther King, Jr., and Malcolm X to exhibit personality traits that were consistent with the 15 characteristics of Maslow's self-actualizing person. This is not unexpected, because Maslow noted that self-actualizing people are likely to be the great reformers of society and the most effective fighters against injustice and inequality, as well as fighters for excellence, effectiveness, and competence.

Evaluative Comments of Maslow's Theory of Needs

Much of Maslow's writing is philosophical or theoretical; he conducted very little research on his need theory or theoretical concepts. Maslow was more of a philosopher of science and a theoretician. Nevertheless, his basic need theory is highly heuristic in terms of generating much scholarly interest and discussion in psychology as well as interest from the general public, primarily because it addresses humanistic concepts that are of great interest to people (e.g., needs in general, physiological survival, safety, love, esteem, and growth).

Regarding research or empirical measurement, Maslow's theory of human motivation influenced the

development of the Personal Orientation Inventory (POI), a psychometric instrument that was developed by Everett Shostrom to measure self-actualization and other dimensions of personal growth. The POI has been used in a number of studies related to self-actualization. Regarding the use of Maslow's needs theory in cross-cultural research, Michael Hagerty conducted a study on quality of life across time and countries. The study used Maslow's original five needs and data from each country that reflected level of need fulfillment for its citizenry.

It is difficult to assess Maslow's work because he died unexpectedly while in the process of redefining his theory of basic needs and his concept of self-actualization. Moreover, he often wrote on individual concepts in his papers, lectures, chapters, and articles and did not develop a conceptualized theory that clearly demonstrated the interrelationship of these concepts and assumptions as one theoretical framework. Another challenge in reading Maslow's writing is his tendency to use multiple or interchangeable terms to discuss an idea or phenomenon. For example, he used *metaneeds*, *metavalues*, and *B-values* with similar or identical meaning, and he used *satisfaction* and *gratification* (of needs) interchangeably. He also discussed *needs*, *desires*, and *impulses* in the same breath or seemingly with the same meaning. Nonetheless, Maslow's theory of basic needs has created much interest across academic disciplines, cultures, and countries.

About Maslow

Maslow was born in Brooklyn, New York, on April 1, 1908, and died from a heart attack in Menlo Park, California, on June 8, 1970. For much of his professional career, he was a faculty member at Brooklyn College and Brandeis University. At Brandeis, he served as chairman of the Department of Psychology; moreover, he was president of the American Psychological Association from 1967 to 1968.

Maslow first published his theory of basic needs in 1943. Other discussions of Maslow's theory or hierarchy of basic needs can be found in his *Motivation and Personality* and his *Toward a Psychology of Being*. The year after Maslow's death, his widow, Bertha G. Maslow, in consultation with some of Maslow's colleagues, published a collection of his articles and papers in the book *Farther Reaches of Human Nature*. This book also contains discussion on his hierarchy of basic needs; furthermore, it includes a comprehensive

bibliography of Maslow's publications and important unpublished papers.

As a psychologist, Maslow's most significant contributions were to the fields of humanistic psychology and transpersonal psychology, wherein many authorities recognized him as a leading pioneer, if not a founder, of these movements or forces in psychology. In addition, Maslow's hierarchy of needs has provided implications and applications for education, business management, and religion. It is a psychological theory with multidisciplinary implications and applications across contexts or settings.

Although Maslow is primarily known for his writings on basic needs and self-actualization, his books, articles, lectures, and papers encompass a number of concepts of humanistic and transpersonal psychology. Most of these concepts are related to his theory of basic needs and self-actualization in some way and include topics such as peak experiences, human aggression and destructiveness, human values, growth, transcendence, humanistic education, creativity, religion, and holistic health. Nevertheless, from the broader perspective, his work's theoretical focus is in the areas of human motivation and healthy personality, and his greatest contribution is probably to the development of positive psychology, humanistic psychology, and transpersonal psychology, or what is generally referred to as the *third force* in psychology.

Frederick D. Harper and Michael Guilbault

See also Emotional Development; Maturation; Social Development

Further Readings

Fridlund, A. J., Gleitman, H., & Reisberg, D. (2003). *Psychology* (6th ed.). New York: Norton.

Hagerty, M. R. (1999). Testing Maslow's hierarchy of needs: National quality-of-life across time. *Social Indicators Research*, 46, 249–271.

Harper, F. D. (1970). *Maslow's concept of self-actualization compared with personality characteristics of selected Black American protesters: Martin Luther King, Jr., Malcolm X, and Frederick Douglass.* Unpublished doctoral dissertation, Florida State University, Tallahassee.

Jourard, S. M. (1974). *Healthy personality: An approach from the viewpoint of humanistic psychology.* New York: Macmillan.

Maslow, A. H. (1943). A theory of human motivation. *Psychological Review*, 50, 370–396.

Maslow, A. H. (1954). *Motivation and personality.* New York: Harper & Row.

Maslow, A. H. (1962). *Toward a psychology of being.* Princeton, NJ: Van Nostrand.

Maslow, A. H. (1968). *Toward a psychology of being* (2nd ed.). Princeton, NJ: Van Nostrand.

Maslow, A. H. (1970). *Motivation and personality.* New York: Harper & Row.

Maslow, A. H. (1971). *Farther reaches of human nature.* New York: Viking.

Shostrom, E. (1964). An inventory for the measurement of self-actualization. *Educational and Psychological Measurement*, *24*, 207–218.

Shostrom, E. L. (1977). *Manual for the personal orientation dimensions.* San Diego, CA: EdITS.

Tosi, D. J., & Lindamood, C. A. (1975). The measurement of self-actualization: A critical review of the Personal Orientation Inventory. *Journal of Personality Assessment*, *39*(3), 215–224.

Maturation

There are many different beliefs about what the dominant forces are in the process of human development. One of these forces is *maturation*, a biological process in which developmental changes are controlled by internal factors. Behaviors that result from maturation, such as walking or secondary sex changes at puberty, are characteristic of the species and are never the result of specific practice or exercise—that is, they are not learned.

Maturation is usually aligned with a belief that heredity is a dominant influence upon development, and the most popular developmental theorist to put forth this belief is Arnold Gesell. Gesell was influenced by G. Stanley Hall's interest in recapitulation theory, which held that the development of the individual recapitulates (or repeats) the evolutionary history of the species' development. Much of recapitulation theory as well as Gesell's early work were influenced by Darwin's theory of evolution as presented in *On the Origin of Species*, published in 1859. Even though Gesell was most influential almost a century ago, his study of education, medicine, and psychology and what he wrote for scientific, professional, and popular audiences continues to have widespread inpact.

Gesell's theory places heavy emphasis on biological forces that provide both the impetus and the direction for development (what we call maturation). Simply put, one of Gesell's basic ideas (based in part on earlier ideas by G. E. Coghill) is that physical structure must be present and developed before function can occur, and behavior is simply not possible if the necessary structures have not yet developed. For example, children cannot walk until they have the structural equipment to do so (including the maturational development of certain muscles as well as the neural organization).

Gesell collected a huge amount of data and maintained that development progresses through an orderly sequence and that the sequence is determined by the biological and evolutionary history of the species. The rate at which any child progresses through the sequence, however, is individually determined by the child's own heredity background (or genotype). Although the rate of development can be artificially altered, it cannot be fundamentally changed. On the other hand, the environment can temporarily affect the rate at which the child develops.

Principles of Development

Gesell sought to unite the basic principles of underlying structural growth with behavioral growth in showing how "psychological growth, like somatic growth, is a morphogenetic process." He described five basic principles of development because they represent developmental principles that occur on a psychological level as well as on a structural one.

1. The Principle of Developmental Direction. The principle of developmental direction assumes that development is not random but proceeds in an ordered fashion. The fact that development systematically proceeds from the head to the toes is a good example of how at any point a developmental trend will be more advanced in the head area than in the foot area.

2. The Principle of Reciprocal Interweaving. This second general principle is modeled after the physiological principle that inhibition and excitation of different muscles operate in a complementary fashion to produce efficient movement.

3. The Principle of Functional Asymmetry. This principle assumes that a behavior goes through a period of asymmetric or unbalanced development to enable the organism to achieve a measure of maturity at a later stage.

4. The Principle of Individuating Maturation. Gesell believed that development is viewed as a process of sequential patterning wherein the patterning is

predetermined and revealed as the organism matures. The principle of individuating maturation stresses the importance of a growth matrix as an internal mechanism that establishes the direction and pattern of development of the individual.

5. The Principle of Self-Regulatory Fluctuation. Finally, the principle of self-regulatory fluctuation proposed that developmental progress is like a seesaw that fluctuates between periods of stability and instability, active growth and consolidation. These progressive fluctuations, as part of a give-and-take much like the principle of reciprocal interweaving, culminate in a set of stable responses.

Individual Differences

Each of these five principles is considered to be characteristic of every child's growth pattern, yet Gesell emphasizes the importance of wide and stable individual differences as well. One of Gesell's unique contributions to the field of developmental psychology was the development of a new method of studying development using moving pictures. Gesell and his colleagues first studied five infants during the first year of life, and each child was rated on 15 behavior traits, such as energy output, social responsiveness, and self-dependence, and then each child was rank-ordered within the group of five on each trait. The same process was repeated 5 years later. A comparison of ranks between the first year and the fifth year showed a remarkable similarity between the two observations on traits such as laterality (handedness), self-dependence, sense of humor, and emotional maladjustment. These findings indicate a certain degree of stability in the development of individual differences, a characteristic that Gesell believed has its source in some kind of biological mechanism.

In another study of 33 infants from infancy through the teenage years, a remarkable degree of stability was also found in what Gesell refers to as trends in mental development. Gesell described individual differences in behavioral development in four different areas: motor behavior, adaptive behavior, language behavior, and personal social behavior.

In each of these four domains, the five general principles of development described earlier continue to function, with individual differences in the rate of development dominating. The child is an integral whole within which the four domains interact with one another while under the control of biological forces expressed through the five principles of development discussed earlier. To summarize Gesell's theoretical outlook, the development of the child is controlled entirely by biologically determined principles of development that produce an invariant sequence of maturation. In turn, this maturational process makes behavioral expression possible. Although individual children progress at their own rate (and the rate is not directly amenable to environmental manipulation), the sequence of development is the same for every child.

Neil J. Salkind

See also Emotional Development; Learning; Moral Development

Further Readings

Ames, L. B. (1989). *Arnold Gesell: Themes of his work.* New York: Human Sciences Press.

Benefice, E., Garnier, D., & Ndiaye, G. (2001). Assessment of physical activity among rural Senegalese adolescent girls: Influence of age, sexual maturation, and body composition. *Journal of Adolescent Health, 28*(4), 319–327.

Gifford, R. (2007). Environmental psychology and sustainable development: Expansion, maturation, and challenges. *Journal of Social Issues, 63*(1), 199–212.

Salkind, N. (2004). *An introduction to theories of human development.* Thousand Oaks, CA: Sage.

Siebenbruner, J., Zimmer Gembeck, M. J., & Egeland, B. (2007). Sexual partners and contraceptive use: A 16 year prospective study predicting abstinence and risk behavior. *Journal of Research on Adolescence, 17*(1), 179–206.

MEAN

The mean is the most often-used measure of central tendency and is usually defined as the sum of all the scores in a data set divided by the number of observations. It can also be defined as the point about which the sum of the deviations is equal to zero.

The formula for the computation of the mean is as follows:

$$\overline{X} = \frac{\Sigma X}{n}$$

where

- The letter *X* with a line above it (also called "*X* bar") is the mean value of the group of scores or the mean.

- The Σ, or the Greek letter sigma, is the summation sign, which directs you to add together what follows it.
- The *X* is each individual score in the group of scores.
- *n* is the size of the sample from which you are computing the mean.

For example, the following data set in Table 1 consists of 30 cases with two variables, reaction time and accuracy, two measures often used in research on cognition and learning.

Table 1   Participant Reaction Time and Accuracy

Participant	Reaction Time	Accuracy
1	7	78
2	5	76
3	8	79
4	11	89
5	7	88
6	6	86
7	13	79
8	8	98
9	9	73
10	10	77
11	8	65
12	7	69
13	6	87
14	7	90
15	8	99
16	9	91
17	9	47
18	16	69
19	7	87
20	12	68
21	4	87
22	7	87
23	6	75
24	14	49
25	11	88

To compute the mean, follow these steps:

1. List the entire set of values in one or more columns such as you see in Table 1. These are all the *X*s.
2. Compute the sum or total of all the values.
3. Divide the total or sum by the number of values.

Applying the formula you see above to the sample data results in the following two means:

$$X_{RT} = \frac{215}{25} = 8.60$$

$$X_{Accuracy} = \frac{1981}{25} = 79.24$$

The mean is sometimes represented by the letter *M* and is also called the typical, average, or most central score.

One should keep in mind the following about the mean:

- The sample mean is the measure of central tendency that most accurately reflects the population mean.
- The mean is like the fulcrum on a seesaw. It is the centermost point where all the values on one side of the mean are equal in weight to all the values on the other side of the mean. That's why the sum of the deviations about the mean always must equal 0.
- The mean is very sensitive to extreme scores. An extreme score can pull the mean in one or the other direction and make it less representative of the set of scores and less useful as a measure of central tendency. This is the argument for using the median as a measure of central tendency.

Neil J. Salkind

See also Median; Mode

Further Readings

Salkind, N. (2004). *Statistics for people who (think they) hate statistics* (2nd ed.). Thousand Oaks, CA: Sage.

MEASUREMENT

Measurement is the assignment of numbers, according to rules, to physical or mental objects, attributes, traits, constructs, or concepts. The purpose of these

numerical assignments is to enable comparisons, assessments, judgments, and evaluations through various mathematical computations and manipulations. These numbers may be obtained via the use of measurement tools or self-reports, or may be based on direct observations or judgments of overt phenomena or behaviors by human observers. Because of the convenience of operations of numbers and the general consensus regarding standard mathematical and statistical rules for comparisons and computations of numbers, measurement is considered a fundamental part of science. Galileo stated that, for the sake of knowledge, we must measure what is measurable and make measurable what cannot be measured.

The measurement of observable attributes of physical objects or phenomena is relatively direct. Human beings have engaged in these physical measurement activities for at least 5,000 years. The earliest known measurement of physical attributes was the use of the Royal Egyptian Cubit, which was defined as the length of the forearm from the elbow to the tip of the outstretched middle finger and is equal to a little less than 21 of today's inches, to measure the physical attribute of length. It is known that the Khufu Pyramid, which was completed around 2750 B.C.E., was constructed based on the use of the Royal Egyptian Cubit. As well, a number of lunar calendars were developed by various ancient civilizations to measure (i.e., to assign numerical values to) the concept of "passage of time."

The measurement of educational and psychological attributes, including covert traits and overt behaviors, however, is a relatively recent phenomenon. One of the earliest rudimentary attempts to assign quantitative descriptors to educational and psychological attributes can be found in the civil service examination system of ancient imperial China. The system, called the *Keju* system, lasted from 606 to 1905. It had been an unofficial practice during approximately the past seven centuries of the system that, when an exam grader encountered an excellent sentence or choice of words in an exam essay, the grader would mark it with one or more circles or dots. Hence, an excellent essay would have a high density of circles and dots and a poor essay would have few to none. This practice did not have the numerical precision of modern educational/psychological measurement, but it was an early method in the use of numerical quantity (i.e., number or visual density of circles and dots) to represent an educational/psychological quality or attribute (i.e., excellent writing ability).

Precise methods of measurement of educational/psychological attributes did not develop systematically until the late 19th century. Gustav Fechner developed the first systematic method of modern psychological measurement when he attempted to measure the intensity of human sensations in 1860. He accomplished this by assigning a numerical degree of intensity based on the concept of "just noticeable difference." The score for a sensation would move up one point on the scale for every "just noticeable difference" in the intensity of a particular sensation, such as pressure. In the 1880s and 1890s, there was a great deal of consensus among the early founding fathers of the then-new discipline of psychology, notably Sir Francis Galton, Charles Spearman, and James Cattell, that measurement was to be a critical characteristic of psychology if this new discipline was to attain the status of a science. These theorists focused primarily on the use of existing measures of physical attributes to gauge psychological attributes such as a person's reaction time, the amount of pressure that induces pain, the amount of time it takes a person to name a color, or Fechner's "just noticeable difference" in sensation—all in a general area of study called psychophysics or anthropometrics. In contrast, Alfred Binet developed the numerical concept of mental age as a measure of the covert psychological construct of intelligence in 1905, when he developed the Binet-Simon Scale of Intelligence. Lewis Terman standardized Binet-Simon's mental age in 1916 by dividing mental age by chronological age and multiplying the result by 100. The result is known as the Intelligence Quotient or Ratio IQ score. Subsequent to these early developments, a variety of new issues related to the choice of scoring metric, the precision or stability of the scores, and the meaningfulness of interpretations and uses of these scores arose; and many new developments took place in the next century. These may be divided into three major areas of concern: scaling methods, reliability theories, and validation methods.

Scaling

Measurement is the assignment of numbers to objects according to rules, and scaling methods provide the rules. The objects of educational and psychological measurement are overwhelmingly covert psychological constructs. These constructs can be divided into those in the affective domain (e.g., feeling, attitude,

sentiment, opinion) and those in the cognitive domain (e.g., achievement, aptitude, ability, knowledge). Whereas cognitive constructs may manifest themselves in the form of correct versus incorrect responses, affective constructs do not generally have correct or incorrect responses. As such, scaling methods differ between the two domains.

Scaling methods in the affective domain were developed by extending earlier psychophysical measures. In 1927, Louis Thurstone developed three scaling methods to replace earlier psychophysical methods. These were the paired-comparisons, the successive intervals, and the equal-appearing intervals methods. With these methods, respondents were presented with statements concerning a certain issue or matter and were asked to respond in different ways, depending on the method chosen. Their responses were calibrated through a series of statistical procedures to yield a numerical score representing the sentiment, opinion, or attitude of each respondent toward that issue. One problem with these methods was the uncertainty as to whether all responses had the same meaning; that is, do they reflect the same psychological dimension?

In 1940, Louis Guttman developed the method of scalogram analysis, better known as the Guttman scale. Through this method, it is possible to ensure that all questions/stimuli are on the same single dimension. Both Thurstone's and Guttman's methods were rather cumbersome, but both are applicable to the scaling of all types of affects. In 1930, Rensis Likert developed an alternative, much simpler method called summated rating, better known today as the Likert scale, specifically for the scaling of attitudes. Many different statements on the same dimension are presented to a respondent, and the respondent is asked to indicate his or her degree of agreement/disagreement or approval/disapproval of each statement. A value is assigned to each response and summed across all statements to form the score for a respondent. This method has become very popular, and variations of this method, generally called Likert-type scales, have been used for a large range of affective measures, ranging from measuring attitudes to judgments of quality to rating of performances. A somewhat less popular method is the semantic differential scaling method developed by Charles Osgood in 1957. This method is particularly appropriate to measure the connotative meaning of concepts in the minds of respondents. A concept is presented and the respondent is to indicate his or her sentiment toward that concept by selecting his or her numerical position between two

bipolar adjectives for each of a number of pairs of such adjectives. The sum of all numerical positions across all adjective pairs is the score for the respondent.

In the cognitive domain, many different rules have been used. For multiple-choice tests, a common rule is to assign a score of 1 to each correct response and a score of 0 to an incorrect response. For essay or performance assessment, analytic or holistic rating scales are frequently used, and ratings are made on the basis of agreed-upon rubrics. Other rules, such as points or credits earned or weighted ratings, have been used as well.

The scores obtained from any of these affective or cognitive scaling methods are generically referred to as observed raw scores. For multiple-choice tests, the raw scores are sometimes transformed to formula scores, which are used to correct inflation of scores due to guessing. For Likert or Likert-type scales in the affective domain and virtually all measurements in the cognitive domain, an alternative to the raw score is the *ability* score. The ability score in theory measures the covert construct directly based on a mathematical relationship between the observed pattern of responses and the covert construct. The statistical processes to derive these ability scores are a group of statistical models and procedures called the Item Response Theory. Some of the most popular statistical Item Response Theory models used today include the Rasch model, the 3-parameter logistic model, the faceted Rasch model, and the multidimensional item response theory model.

When determining the appropriate scaling method, one important question is what type of numbers should be used to calibrate a particular object. The answer lies in the kind of mathematical or statistical analysis or comparisons one wishes to be able to perform with the numbers resulting from the scaling process. In 1959, S. Smith Stevens identified a number of ways in which numbers are used to represent objects. He summarized these into four major "scales of measurement," which he called nominal, ordinal, interval, and ratio scales. Nominal scales are numbers that serve as labels, such as numbers on football jerseys. Only a limited array of operations, such as judgments of equality, frequency count, mode, and chi-square tests, is sensible. Ordinal scales are numbers with an order, such as street numbers or rankings. The difference between adjacent numbers in terms of the value of what these numbers represent is not constant. Greater-than or less-than comparisons can be made with these numbers. One can also meaningfully

compute the median, percentile scores, and Spearman rho, or perform sign tests or run tests based on these numbers. Interval scales are ordinal numbers with equal interval values between adjacent numbers. Examples include calendar years and temperature. Sensible computations include additions, subtractions, arithmetic means, standard deviations, correlations, *t*-test, and *F* test, among others. Finally, ratio scales are interval numbers with an absolute zero value, such as the absolute zero point in Kelvin temperature (where there is no molecular activity). All mathematical/statistical computations can be sensibly performed on these numbers. In general, scores that are at the nominal or ordinal level can be meaningfully analyzed only through nonparametric statistical methods, whereas those at the interval and ratio level can use more powerful parametric methods. However, some more recent methods, such as log-linear modeling or logistic regression, have enabled the analyses of nominal or ordinal data via parametric statistics by transforming these data onto an interval or ratio metric. Additionally, conventional parametric statistics are known to be robust such that their uses on ordinal data generally lead to only small errors.

For the measurement of covert psychological traits or constructs, a desired goal of measurement is to attain scores at the interval or ratio level. Raw scores in common educational and psychological measurements are generally considered to be at the ordinal level. However, ability scores obtained through Item Response Theory approaches are considered to be at the interval level. An alternative is to transform raw scores to standard scores; the latter are considered to be at the interval level of measurement.

Standard scores are results of particular linear or nonlinear transformations of raw scores to particular scoring metrics. Although it is possible to compare and to perform statistical operations with raw scores, a single raw score by itself has no inherent meaning. To interpret a single raw score, it is necessary to compare that raw score against something else. One of the most common practices is to compare a person's raw score against the raw scores of that person's peers. This approach is called *norm-referenced interpretation.* Another common practice is to compare the raw score against some agreed-upon absolute standard (e.g., passing score). This is called *criterion-referenced interpretation.* In order to facilitate individual score interpretation, raw scores are usually transformed to a different metric such that norm-referenced or criterion-referenced information is built into the new scoring metric. Examples of these transformed scores include percent-correct, ranks, percentiles, and grade equivalent scores. Standard scores, which include linear standard scores and normalized standard scores, are a category of scores transformed from raw scores with built-in norm-referenced information. Some of the most commonly used standard scores include *z* scores, *T* scores, stanines, deviation IQ scores, and NCE scores.

Reliability

Measurement is fundamentally a sampling process. The score obtained for a given individual through any measurement procedure can represent the affective responses or cognitive performance of that individual only at that moment under that particular set of circumstances. As such, the observed score represents only the result of a single sampling probe. A variety of random and systematic situational factors can affect the observed raw or standard score, as well as the pattern of response used to derive a person's ability score. Therefore, part of the observed score or the ability score reflects unintended extraneous factors other than the true score of the individual. The part of the score that is the result of these unintended extraneous factors is called *measurement error.*

Measurement error is gauged through one or both of two indices: reliability coefficient and standard error of measurement. Reliability coefficient is technically defined as the proportion of variation in observed scores that is attributable to variation in true scores. As such, the more measurement error there is in the observed scores, the lower the value of the reliability coefficient, which ranges between 0 and 1. The higher the value of the reliability coefficient of a set of observed scores, the fewer measurement errors are contained in the observed scores and the more accurately and consistently these observed scores represent their corresponding true scores.

The estimation of the reliability coefficient value for a set of observed scores is accomplished through one of a number of alternative data collection and analysis strategies based on statistical measurement theories of relationships between observed and true scores. These strategies are used under the condition that the data collected meet certain statistical assumptions that underlie the measurement theory on which the particular strategy is based. The most commonly used assumptions are the Parallel Tests Assumptions

within the Classical Theory, which require the collection of data from two interchangeable tests that meet certain statistical requirements. Under these assumptions, reliability coefficients can be estimated via the test-retest method, the equivalent forms method, or the split-half method. Another very popular approach is the Cronbach alpha method used under the Essentially Tau-equivalent Assumptions within the Classical Theory. These assumptions are similar to the Parallel Tests Assumptions but are less restrictive in terms of requirements of data characteristics. Yet another popular approach is the Kuder-Richardson Formula-20 (KR-20) method, which is a special case of Cronbach alpha when responses to each question/item are scored dichotomously as either 0 or 1.

In contrast to the Classical Theory, the Generalizability Theory is based on the Randomly Parallel Tests Assumptions—or alternatively called the Domain Sampling Model—which only require that the measurement data be collected from a random sample of respondents, items, raters, and so on, from the universe of similar respondents, items, or raters. Under the Generalizability Theory, there is theoretically a unique reliability coefficient for each unique interpretation, generalization, and usage of the same measurement instrument, procedure, system, or protocol. Each of these reliability coefficients is called a G-coefficient. Finally, the Item Response Theory approach assumes that all items measure a single dimension and that items are mutually and locally independent. This approach does not produce a reliability coefficient per se. Instead, it generates item and test information functions, which are a conceptual analog to reliability coefficients.

Another indicator that is often used in addition to reliability coefficients is the standard error of measurement. This is an indicator of the degree of random fluctuation of scores that can be expected to be found in the observed scores obtained via a particular measurement procedure using a particular instrument. It is used to determine the size of the margin of error around a given observed score. The larger the standard error of measurement for a particular procedure, the less precise are the observed scores and the more random errors are contained in the observed scores from that procedure. One commonly used margin of error is the 95% confidence interval, which is a range of scores within which there is a 95% chance that the true score of a person will lie. The upper end of this range is obtained by adding 1.96 times the standard error of measurement value to the observed score, whereas the lower end is attained by subtracting from the observed score 1.96 times the standard error of measurement.

Validity

In educational and psychological measurement, the observed score is typically based on the measurement of an observable indicator of some unobserved theoretical construct. The unobserved construct is of interest, although it is the observable behavior or phenomenon that is being measured. For example, the scores based on observable responses to a number of questions might be used as an indicator of the theoretical construct of self-efficacy, or the frequency of a certain observable behavior is used as an indication of a child's unobservable construct of motivation. A major concern is whether the inferential leap from the observed score to the unobserved construct is justified.

Additionally, educational and psychological measurements are rarely taken just for the sake of measuring a certain unobserved construct. Rather, the scores usually form the basis for decisions and actions. For example, the scores on an intelligence test may be used to decide whether to place a child in a gifted education class; or the score on a behavioral checklist is used to determine whether a client should receive psychotherapy; or the score on a vocational aptitude test is used to determine whether to hire a prospective employee. Even if the inference from the observed to the unobserved construct is justified, there is no guarantee that the way the score is used to make decisions about an individual is appropriate. For example, there may be evidence to support the interpretation of the scores on a personality test as indicators of the unobserved construct of personality, but that alone does not justify using those scores as part of a particular employment decision.

Validity is the adequacy and cogency of evidence to support score-based inferences and decisions. In the literature, there are a variety of opinions as to what constitutes validity and what is required for the claim of validity among researchers. Since the 1960s, educational psychologists have been abiding by the opinions of the Joint Committee on Standards for Testing as the authoritative guide for validation and other measurement issues. The Joint Committee is a 16-member professional committee representing the American Educational Research Association, the

American Psychological Association, and the National Council on Measurement in Education, and it is charged with the task of setting professional standards of quality for educational and psychological testing and measurements. The standards produced by this committee are also endorsed and observed by a number of other professional organizations. Because of the rapidly changing nature of measurement and assessment in the past half-century, a new Joint Committee was formed approximately every decade to update and revise as needed the professional standards of quality of measurement. The latest set of standards was published in 1999. According to the 1999 *Standards*, validity is the degree to which evidence and theory support the interpretations of test scores entailed by proposed uses of tests.

Evidence to support validity may be gathered from a variety of sources and in a variety of formats. Some types of evidence may be empirical in nature, others may be based on literature reviews, and yet others may be based on logical analyses. The 1999 *Standards* further suggests five possible sources of validity evidence: process, consequences, internal structure, test content, and relations with other variables. Evidence based on process refers to evidence that shows that the physical or mental processes that a respondent or test-taker goes through while performing the measurement tasks authentically correspond to the theoretical processes suggested by the intended construct. An issue of primary concern for this source of evidence is authenticity of the measurement process. Evidence based on consequences refers to evidence that the use of a measurement procedure has achieved its intended purpose and has minimal, if any, unintended negative consequences on the individual, the system, or society at large. Some concepts within this type of evidence include systemic validity, washback, and instructional validity. Evidence based on internal structure refers to evidence that the content components of a test or measurement procedure and their interrelationships are consistent with the theoretical structure of these components or factors within the intended construct. This type of evidence is sometimes called *structural validity* and is most often gathered via factor analysis techniques.

Evidence based on content refers to evidence that shows that the contents of a test or measurement tool such as the questions asked or tasks to be performed form a representative and relevant sample of the larger domain of possible questions and tasks for the intended construct. This source of evidence had been referred to as content validity or content-related evidence in earlier *Standards* and literature. Some issues of concern within this type of evidence include curricular validity, face validity, teaching to the test, and opportunity to learn. Evidence based on relations with other variables refers to evidence that the scores correlate to scores on other variables as expected by the theory entailed by their interpretation and use. This source of evidence is primarily statistical in nature and includes a large array of possible approaches and techniques. Some of the better-known approaches and techniques include criterion-related evidence, predictive validity, concurrent validity, nomological validity, hit rate, sensitivity, multitrait-multimethod matrix, incremental validity, and validity generalization.

The exact mix and amount of evidence that is needed to support a claim of validity differs from one measurement situation to another and from one use or purpose to another, depending on feasibility and what may provide the most compelling set of arguments for the particular interpretation or the particular use. If many pieces of strong evidence are not available, a few pieces of cogent evidence are preferred over many pieces of feeble evidence.

Two issues that cut across all areas of validity concerns are fairness and special accommodations. Both are legally mandated and are potential causes of litigation. Fairness refers to the absence of measurement bias, which exists when deficiencies in the measurement procedure have led to different meanings of scores for people in different groups. The groups of concern are usually based on race, gender, ethnicity, language, or other sociopolitical variables. Whenever feasible, steps should be taken to detect and eliminate test bias prior to the use of measurement procedures. The primary method to detect test bias today is through the detection and elimination or modification of biased items or tasks within the measurement tool. This is accomplished through a two-step process of differential item functioning, better known by its acronym of DIF, which consists of analyses followed by sensitivity reviews. DIF analyses are a group of statistical techniques used to identify items or tasks that have somehow led to different responses from respondents who have the same level of trait or ability but are members of different groups. Sensitivity review is a judgmental exercise in which an independent panel of content area experts from different groups reviews the items that have shown DIF. The purpose is to

discern whether bias is the cause of DIF for each of the items. Biased items would then be eliminated or modified as appropriate.

The phrase *special accommodations* refers to the need to modify the measurement procedure or setting in order to attain fairness for respondents or test-takers with known mental, learning, or physical disabilities. Some examples of special accommodations include providing extra time for individuals with learning disabilities or providing a reader for a blind test-taker. The exact effects of special accommodations on validity are not well understood today. In general, without special accommodations, the validity of interpreting and using scores will suffer for individuals with disabilities. Yet the use of special accommodations may introduce extraneous factors that may adversely affect validity. The trick is to find the most appropriate accommodation for an individual to ensure fairness on one hand and to avoid introducing construct-irrelevant factors into the score on the other.

Hoi K. Suen

See also Aptitude Tests; Criterion-Referenced Testing; High-Stakes Testing; Intelligence Tests; Mental Age; Multiple-Choice Tests; Norm-Referenced Tests; Personality Tests; Standardized Tests; Standard Scores; *T* Scores; Testing

Further Readings

American Educational Research Association. (1999). *Standards for educational and psychological testing.* Washington, DC: Author.

Linn, R. L. (Ed.). (1989). *Educational measurement* (3rd ed.). New York: Macmillan.

McDonald, R. P. (1999). *Test theory: A unified treatment.* Mahwah, NJ: Lawrence Erlbaum.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

MEASUREMENT OF COGNITIVE DEVELOPMENT

Very few would doubt that the most influential account of cognitive development over the previous century has been that of Piaget and his Genevan school of researchers. Piaget's cognitive developmental theory is an empirical part of his larger philosophical theory of genetic epistemology. Through investigations involving his own children during their infancy and the problem solving of hundreds of thousands of Swiss schoolchildren over six decades, he proposed a hierarchical sequence of four cognitive developmental stages: sensori-motor, pre-operational, concrete operational, and formal operational stages. In the post-*Sputnik* era, U.S. researchers, in particular, turned to cognitive developmental theory to inform classroom practice and curriculum design. Concerted attempts were made to augment the prevailing psychometric accounts of IQ by investigating the quantification of particular cognitive developmental concepts such as conservation, or, more generally, of Piaget's stages. In educational psychology, the term *measurement* is used interchangeably with *quantification* and *statistics*, but in this entry, the term is restricted to describe just those scales with iterative units.

It is somewhat ironic that the cognitive developmental theory of Piaget was dismissed in the United States and United Kingdom because, in part, it did not yield the quantitative indicators expected by psychologists of that era. His theory was decidedly qualitative, in spite of its logico-mathematical structure, at a time when quantitative approaches dominated the field of educational psychology. Decades later, when psychology is more accepting of qualitative theories and the principles for measuring Piagetian cognitive development are now quite well established, some mainstream educational psychologists seem to regard his theory as passé. Those who have persisted in the attempt to measure cognitive development have managed to satisfy the most stringent requirements for measurement in the field of psychometrics—those of the Rasch model for measurement.

Psychometricizing Piaget

The conventional view of the state of play in the 1960s can be summarized by looking to the book *Measurement and Piaget* and a review of it that appeared in the year following its publication. *Measurement and Piaget* recorded the proceedings of a Conference on Ordinal Scales of Cognitive Development sponsored by CTB/McGraw-Hill in 1969. In an interesting review of the book devoted to the proceedings of a conference in 1969, Wohlwill summarized rather succinctly what has been regarded as the lack of interest by Piagetians in quantification of

cognitive development where the title of the book *Measurement and Piaget* was juxtaposed against his review title "And Never the Twain Did Meet." He then painted a picture of the archetypal Piaget deliberately relegating the problems of "psychometricizing" Piagetian measures to a province beyond his concern, on the grounds that Piaget asserted that he had "no interest whatever in the individual." More tellingly, he then described other developmentalists at the conference as quite ready to let the topic of measurement lie in the limbo to which Le Patron had relegated it.

It would be too easy, as well as erroneous, to disregard Piaget's views on psychological measurement as being ill informed. In his discussion on the role and techniques of psychological research, *The Place of the Sciences of Man in the System of Sciences*, the critique that Piaget makes of the psychometric techniques of the time shows him to be much better informed than many of his contemporaries. It is only now, after the end of that century, that one might fully comprehend Piaget's prescience: His claim was that the chief difficulty with the sciences of man, and indeed with all of the life sciences, lies in the absence of units of measurement, the property that was common to the measurement systems that abound in the physical sciences. This bold and informed assertion was published in English, after the 1969 measurement conference referred to above, but in print before both the *Measurement and Piaget* text of 1971 and the "Never the Twain Did Meet" critique of 1972.

Considerable work on the quantification of cognitive development took place in Geneva and in France following directly from the work of Piaget. The research of Longeot in France had impacts outside of the French-speaking world; unfortunately, the progress of Nassefat and Uzan in Geneva went largely unnoticed by Anglophones. However, the more recent successes in creating genuine measurements of cognitive development are underlined by the use of the Rasch model—the only widely available technique for constructing unidimensional, interval-level measurement scales from what must be regarded as ordinal-level qualitative responses collected from a Piagetian perspective. The early French-language research had opted for Guttman scalogram analysis and Loevinger's H statistic over factor analytic and other techniques based on correlational principles. Nowadays, Rasch measurement is touted as the solution for the measurement of cognitive development because it instantiates the iterative units of measurement that

Piaget had argued were missing from psychometrics routinely used in the 1960s and 1970s.

Measurement Scales for Cognitive Development

The Uzgiris-Hunt Ordinal Scales of Psychological Development are the widely used standardized indicators of cognitive development based on the theories of Piaget. The subscales focus on the abilities identified as key components of Piagetian sensori-motor development (e.g., Object Permanence, Means-Ends, Operational Causality). In terms of the context of this entry, the measurement of cognitive development, it is worth noting that the explicitly ordinal nature of cognitive developmental data is clearly asserted both in the title of the Uzgiris-Hunt scales and in their source text.

It would appear that one of the persistent problems for researchers wanting to quantify cognitive development is the obvious mismatch between easily scored standardized testing formats and the flexible, unstructured interview procedures that formed the backbone of the Genevan research. Results obtained by such a flexible procedure as the *méthode clinique* do not lend themselves easily to statistical treatment. Indeed, even Piaget himself subscribed to that view. As a consequence, standardized individual interview procedures were developed, class tasks were substituted, or pencil-and-paper tests were written to provide the sort of data amenable to statistical analyses. Two of Piaget's assistants worked directly on the standardization of the Genevan data collection procedures: Vinh-Bang, who continued to work with Piaget in Geneva, never published his results; Noelting set up his own research team in Quebec and adopted quantitative methods as the basis for his research into cognitive development.

The research conducted by Shayer's team in the United Kingdom based on specially developed class tasks of concrete and formal operational tasks (e.g., conservation, pendulum) yielded the only nationwide summary of cognitive development anywhere in the world, providing results remarkably in accord with the Genevan treatise on the transition from concrete to formal operational thinking published in the 1950s. In the United Kingdom, this led to a revelation of the mismatch between the profiles of adolescent thinking capacities and curriculum demand in secondary schools, and successful attempts to maximize cognitive development with science-based classroom interventions.

Conclusion

Rasch measurement procedures have been used to confirm the unidimensional nature of interval-level measures created by standardized procedures such as the class-task and multiple-choice tests. Data from preoperational, concrete operational, and formal operational solutions to traditional Piagetian qualitative interviews, including many of the conservations and the pendulum, can now be scored using schedules derived directly from the Genevan sources and converted to interval-level measures of student ability and task difficulty, thereby quantifying cognitive development over time and revealing the relationships between cognitive development and school achievement.

Trevor G. Bond

See also Cognitive Development and School Readiness; Piaget's Theory of Cognitive Development

Further Readings

Bart, W. M., & Airasian, P. W. (1974). Determination of the ordering among seven Piagetian tasks by an ordering theoretic method. *Journal of Educational Psychology*, 66(2), 277–284.

Bond, T. G. (1995). Piaget and measurement II: Empirical validation of the Piagetian model. *Archives de Psychologie*, 63, 155–185.

Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.

Goldschmid, M. L., & Bentler, P. M. (1968). The dimensions and measurement of conservation. *Child Development*, 39(3), 787–802.

Green, D. R., Ford, M. P., & Flamer, G. B. (Eds.). (1971). Measurement and Piaget. *Proceedings of the CTB/ McGraw-Hill Conference on Ordinal Scales of Cognitive Development.* New York: McGraw-Hill.

Reuchlin, M. (1968). Measurement in psychology. In P. Fraisse, J. Piaget, & M. Reuchlin, *Experimental psychology: Its scope and method* (Vol. 1). London: Routledge & Kegan Paul.

Shayer, M., Küchemann, D. E., & Wylam, H. (1976). The distribution of Piagetian stages of thinking in British middle and secondary school children. *British Journal of Educational Psychology*, 46, 164–173.

MEDIA LITERACY

Media literacy is the process of critically analyzing media messages and the ability to compose messages using media tools and technologies. Media literacy is defined as an extended conceptualization of literacy, the ability to access, analyze, evaluate, and communicate messages in a wide variety of forms. The term *access* generally means the ability to locate information or find messages and to be able to comprehend and interpret a message's meaning. The term *analysis* refers to the process of recognizing and examining the author's purpose, target audience, construction techniques, symbol systems, and technologies used to construct the message. The concept of analysis also includes the ability to appreciate the political, economic, social, and historical context in which media messages are produced and circulated as part of a cultural system. *Evaluation* refers to the process of assessing the veracity, authenticity, creativity, or other qualities of a media message, making judgments about a message's worth or value. Finally, the definition of media literacy includes the ability to *communicate* messages in a wide variety of forms (using language, photography, video, online media, etc.). Media literacy emphasizes the ability to use production processes to compose and create messages using various symbol systems and technology tools. In recent years, media literacy has also been described as an expanded conceptualization of literacy, a view that many literacy educators embrace.

Media literacy is primarily conceptualized as a learning outcome within an educational framework that aims to give children and young people opportunities to learn about mass media, popular culture, and communication technologies. *Media literacy education* and *media education* are terms used to refer to the pedagogical processes used to develop media literacy.

Key Concepts of Media Literacy

Because there are many different types of genres and formats within specific media and communication technologies, media literacy programs may address these specific forms directly. For example, media literacy programs have included a focus on critical analysis of newspapers and television news, print and TV advertising, magazines, popular music, contemporary film, and participatory media such as video games and the Internet. Many media literacy advocates and educators make use of a unifying framework: *key concepts* or questions that identify the central ideas associated with media literacy learning. The key concepts can be explored with children of different ages and

with different types of media messages. These include the following:

1. *Messages are constructions.* The media do not present simple reflections of external reality. Rather, media messages are carefully crafted constructions that are the result of many decisions and determining factors.

2. *Media messages are constructed using a creative language with its own rules.* Individual media messages can be recognized within specific genres (like cartoons, news, advertising, romance, horror, biography). Media messages make use of symbol systems and codes and conventions that can be verbal, visual, auditory, musical, narrative, or digital. For example, in narrative films for children, the bumbling or evil adult is a character stereotype that is commonly used in creating conflict.

3. *Audiences actively interpret messages.* People construct meaning as they consume media messages. Message interpretation varies according to individual factors such as developmental level, personal needs and anxieties, situational factors, racial and sexual attitudes, and family and cultural backgrounds.

4. *Media have embedded values and points of view.* Explicitly or implicitly, media express ideological messages about issues such as human nature, social roles, authority and power, and the distribution of resources. Media messages provide the majority of the observations and experiences that people use to develop personal understandings of the world and how it works. Much of people's sense of reality is based on media messages that contain representations that have been specifically constructed to embody points of view, attitudes, and values.

5. *Media have commercial implications and exist within an economic context.* Media literacy aims to encourage an awareness of how the media are influenced by commercial considerations, and how economics and power affect message content, production techniques, and distribution. Many media products that children and young people consume are created as part of global business interests. Questions of ownership and control are important because a relatively small number of individuals decide what others watch, read, and hear in the media.

International Perspectives on Media Literacy

Most people would agree that Great Britain has the most well-established program of media literacy education in the world. Media education first appeared in the 1930s in England from a classical tradition of literary criticism, which established a premise that modern society and its cultural manifestations were alienating and mechanistic. Scholars began moving away from a "high culture/low culture" dichotomy toward analyzing the media through an examination of authorship, audience reception, meaning-making, and cultural identity in a sociocultural context. In 1985, Len Masterman's book *Teaching the Media* became widely influential, reaching an international audience of scholars and educators with interests in media literacy. By 1989, media education was mandated as a compulsory part of British education and located as a subject area within English. Many organizations have been important in supporting the work of teachers and students, including the British Film Institute, the English and Media Centre, the Institute on Education at London University, and many others.

The European Union has also provided support for media education. Education ministers of the European Union recognized the principle of media education as a basic entitlement of every citizen, from the earliest years of schooling. Since then, most European nations have included some kind of requirement for media education in their school curricula. Regulatory bodies have developed an interest in media education as a counterbalance to the increasingly complex problem of media regulation in a digital age, and there has been some corporate interest in sponsoring media literacy activities.

Media literacy has had a long history in Canada, where the Ontario-based Association for Media Literacy has supported the work of educators there since 1978. This organization maintains a Web site; publishes a newsletter; and organizes conferences, summer institutes, and workshops for teachers, and its success has inspired the creation of similar associations in eight other provinces. In Australia, Australian Teachers of Media (ATOM) is an independent, nonprofit, professional association for media teachers and others who wish to use media effectively in their classrooms. The organization sponsors competitions for student-produced multimedia, provides opportunities for teacher education, and hosts statewide conferences.

In most of Asia, media literacy is still in an initial stage of development. Media educators in Japan, Korea, and other Asian nations work primarily in nonschool settings or promote media literacy as a component of parent education. In 2005, a 4-day conference was held in China to introduce the concept of media literacy and to explore how parents and teachers can help teach young people to critically evaluate the media.

Media Literacy Programs

Since the 1970s, media literacy has developed through the work of educators, artists, media professionals, and scholars who have implemented programs in schools and nonschool settings to explore mass media, popular culture, and communication technologies with children and young people. Many factors may lead educators to begin to integrate media analysis and media production activities into the curriculum. Motivations may include a focus on increasing student motivation for learning; responding to ubiquitous elements of media culture, including sexism, violence, and materialism; expanding appreciation for alternative or noncommercial media; reducing the power of U.S. media corporations to control culture; enhancing technology skills; responding to student learning styles; strengthening students' recognition of how print and visual media work as forms of expression and communication; or enabling students to explore the constructed nature of cultural identity, social power, and values. Generally, most media literacy initiatives occur as the result of the initiative of a single individual or small team, working at a local level within the contexts of their school, community, or nonprofit youth-serving organization.

Methods of instruction emphasize the process of critical analysis of a variety of print, visual, electronic, and digital texts through questioning and active discussion, as well as opportunities for children and young people to represent their own ideas through creating media in a wide variety of forms. Using critical questions to stimulate students' active cognitive response is increasingly a common classroom practice, and this instructional strategy has been extended to include the texts of popular culture, including television, movies, magazines, popular music, and participatory media such as video games and the Internet. Other instructional methods include role-playing, simulation, and media comparison-contrast activities. Media literacy has been integrated into all of the K–12 subject areas, and numerous resource materials exist to support the work of elementary and secondary teachers in integrating media literacy into existing instruction. However, the largest number of media literacy programs, curriculum materials, and resources have been developed to align with the subject areas of English language arts and health education.

Media Literacy in English Language Arts

Media literacy has long been part of English language arts education in many K–12 schools in a number of English-speaking nations. In the United States, the National Council of Teachers of English first adopted policy language supporting media literacy education in 1975, stating that the organization should continue

> to encourage teacher education programs which will enable teachers to promote media literacy in students; and cooperate with organizations and individuals representing teachers of journalism, the social sciences, and speech communication to promote the understanding and develop the insights students need to evaluate critically the messages disseminated by the mass media.

In 2003, the organization adopted a policy on composing with nonprint media, encouraging pre-service, in-service, and staff development programs to focus on new literacies, multimedia composition, and a broadened concept of literacy.

Educators with interests in media literacy generally adopt perspectives from the disciplines of the humanities, semiotics, and cultural studies to guide their work, although some make use of media effects or psychological research in learning theory. In the United States, media literacy has also been used in large high schools as a theme to create small learning communities, enabling teachers from several subject areas to make use of media literacy concepts in their classrooms with a smaller group of learners. Research has shown that media literacy education can improve adolescents' reading comprehension, writing, and critical thinking skills.

A school-based media literacy program in a particular school might take many forms. At the middle school or high school levels, some teachers will include popular culture texts and media studies topics into their existing English or social studies classes. Other teachers may implement special instructional units on specific media genres, topics, or themes, such as

journalism and the role of news in society, advertising and cultural identity, or an exploration of stereotyping in film or television. Others will examine similarities and differences between literature and film. Materials and staff development programs for teacher education are available in many parts of the United States and around the world to support the development of teachers' understanding of media literacy.

Media Literacy in Health Education

As a component of health education, media literacy curriculum materials have been created to address the problem of media violence and aggression, nutrition, body image, substance abuse prevention, and other topics. Most health educators in the United States now include some focus on analyzing advertising in the context of understanding substance abuse, including alcohol and tobacco. Hundreds of regional health conferences between 1995 and 2000 featured presentations and workshops demonstrating media literacy as part of health education, and health professionals are a major subgroup of the membership of the Alliance for a Media Literate America (AMLA), one of the two national membership organizations for media literacy. Major federal organizations, including the Centers for Substance Abuse and Prevention (CSAP), the White House Office of National Drug Control Policy (ONDCP), and the National Institute for Child Health and Human Development (NICHD), have supported media literacy as a means to promote child and adolescent health by developing curriculum materials, hosting teacher education conferences, or providing funding for programs.

Research has shown that media literacy education is associated with reduced susceptibility to tobacco use among children and adolescents and increased skepticism about perceptions of the thin ideal in beauty and fashion magazines among adolescent girls. In one study, parents of preschoolers who received media literacy education emphasizing nutrition and food advertising learned how to critically analyze television commercials about food products, which resulted in increased awareness of the need to communicate to their children about what is truthful in media messages.

Youth Media Programs

An increasing number of nonschool programs are using media literacy concepts in their work with children and young people in after-school programs,

summer camps, and other nonschool settings. The term *youth media* is emerging to describe the work of a broad range of nonschool organizations that use a variety of media and technologies to serve youth. Such programs typically involve older children and adolescents in some forms of critical analysis and media production activities. Youth media practitioners may emphasize media production as a form of social activism in local communities, and most youth media programs reflect the particular values of the social justice advocates, youth development specialists, media artists, and technology access providers who conduct these programs. Authentic representation and "voice" are emphasized in programs that are designed to give adolescents opportunities to strengthen leadership skills and advocate for issues of concern to them.

Renee Hobbs

See also Literacy; Social Development

Further Readings

Bazalgette, C., Bevort, E., & Savino, J. (Eds.). (1992). *New directions: Media education worldwide.* London and Paris: British Film Institute & CLEMI.

Buckingham, D. (2004). *Media education: Literacy, learning and contemporary culture.* London: Polity.

Center for Media Literacy. (2005). *Literacy for the 21st century: An overview and orientation guide to media literacy education. Part I: Theory.* Santa Monica, CA: Center for Media Literacy. Retrieved from http://www.medialit.org/pdf/lit2105.pdf

Hobbs, R. (2007). *Reading the media: Media literacy in high school English.* New York: Teachers College Press.

National Alliance for Media Arts and Culture. (2003). *A closer look: Case studies from NAMAC's youth media initiative.* San Francisco: Author.

Primack, B. A., Hobbs, R., Switzer, G. E., Land, S., Fine, M. F., & Gold, M. (2006). Associations between media literacy and adolescent smoking. *Journal of Adolescent Health, 38*(2), 93–94.

Schwarz, G., & Brown, P. (Eds.). (2005). *Media literacy: Transforming curriculum and teaching* [National Society for the Study of Education Yearbook, 104(1)]. Oxford, UK: Blackwell Synergy.

MEDIAN

Quantitative research in educational psychology typically begins with a description of the sample of

subjects. This initial stage of analysis focuses on describing the center, shape, and spread (dispersion) of the observed data points. Various measures are available for describing the center, shape, and spread of a distribution of observed data points. In most instances, the researcher begins by describing the center point or central tendency of the distribution. Measures of central tendency (center) include the mean, median, and mode. The median is an often-used measure of central tendency because of the desirable characteristics it possesses.

The *median* is technically defined as the data point that splits the distribution of data in equal halves. That is, half of the data points fall above the median and half of the data points fall below the median. For example, suppose a mathematics assessment is administered to a group of students and the median score is found to be 74. This indicates that 50% of the students scored better than 74 and 50% scored worse than 74. Thus, the median for a data set is equivalent to the 50th percentile.

Calculating the median is rather straightforward, particularly with a small number of observed data points. When done by hand, the median can be calculated by arranging a set of observed data points in ascending (or descending) order. For example, consider seven hypothetical scores on the aforementioned mathematics assessment:

$$33 \quad 69 \quad 77 \quad 66 \quad 79 \quad 84 \quad 74$$

To calculate the median, the scores are first rearranged in ascending order,

$$33 \quad 66 \quad 69 \quad 74 \quad 77 \quad 79 \quad 84$$

and the middle score is identified. In this case, half of the scores fall above a score of 74 and half fall below a score of 74.

In a distribution consisting of an even number of observed data points, there is no "true" median. Rather, the median is calculated as the average of the middle two scores. Reconsidering the previous example with an additional observation

$$33 \quad 66 \quad 69 \quad 74 \quad 76 \quad 77 \quad 79 \quad 84,$$

the centermost values are 74 and 76, which, when averaged, result in a median value of 75 (150/2).

A particularly desirable characteristic of the median is that it is unaffected by extreme or outlying data points. Unlike the mean, the median does not take into account the value of each data point in its calculation. Thus, the median often provides a more accurate representation of the center of a distribution of data points. To illustrate why this is so, reconsider the following data representing mathematics assessment scores for eight students:

$$33 \quad 66 \quad 69 \quad 74 \quad 76 \quad 77 \quad 79 \quad 84$$

This data set is fairly evenly distributed except for the lowest score of 33. This data point is 41 points lower than the median value of 74 and 33 points lower than the next smallest value of 66. When calculating the average or mean score, all of these data points are taken into consideration. The mean of the data set with the extreme observation is 69.75, whereas the mean without the extreme observation is 75. On the other hand, the median score without the extreme observation does not change drastically. Whereas the median for the data set with the extreme observation is 75, the median without the extreme observation increases slightly to 76.

Unlike the mean and mode, the median can be used to identify the center of both ordinal data and interval data. For example, consider the following distribution of final grades in a statistics class:

$$D \quad C \quad C \quad C \quad B \quad B \quad B \quad A \quad A$$

In this instance, the mean could not be calculated unless the grades were transformed to a numerical scale. Furthermore, even if the grades were transformed to a numerical scale (e.g., F = 0, D = 1, C = 2, B = 3, and A = 4), the mean would not necessarily provide meaningful interpretation about the average letter grade in the class. Likewise, the mode is of limited use in this situation because of the bimodal nature of this distribution. On the other hand, the median clearly indicates that half the students received a letter grade of B or better and half received a letter grade of B or worse.

One approach to identifying the shape of the distribution involves using the relationship between the median and the other two measures of central tendency, namely, the mean and the mode. A particularly desirable characteristic for distributions of data is that they resemble a bell-shaped or normal curve. When this is the case, the median will be equal to the mean and the mode. Conversely, if the median is not equal to either or both of these values, the distribution is

considered to be skewed in a positive or negative direction. Positive or right-skewed distributions occur when the median is lower than the mean. In contrast, negative or left-skewed distributions occur when the median is larger than the mean. In skewed distributions, the mean is pulled in the direction of the extreme value, whereas the median is unaffected by the skewness and remains relatively consistent in terms of identifying the center of the distribution.

In summary, the median is a measure of central tendency that is not affected by extreme values or skewed distributions. Extreme values and/or skewed distributions are present in most data sets, particularly those related to standardized test scores. In such situations, the median can be a useful measure of central tendency either by itself or with the mean and/or mode.

Greg W. Welch and Chris S. Meiers

See also Descriptive Statistics; Measurement; Qualitative Research Methods

Further Readings

Shavelson, R. J. (1996). *Statistical reasoning for the behavioral sciences* (3rd ed.). Boston: Allyn & Bacon.

Memory

The study of human memory is a large enterprise, and research on the topic has applications to the field of education almost too numerable to list. During a normal day, students are asked to learn and retrieve information, remember to complete upcoming assignments, and deal with the frustrating shortcomings of memory familiar to us all. However, in spite of the large body of memory research that has accumulated over the years, several basic principles have been established that, if properly understood, provide a thorough and relatively concise overview of the field. These principles can best be established by crystallizing the vast body of knowledge concerning research on human memory into several basic and empirically supported assertions. The remainder of this entry provides breadth and detail to each of the following assertions with an eye toward the theoretical, conceptual, and empirical underpinnings of each.

First, most memory researchers now agree that there are several distinct types of memory, each with unique properties such as the amount of information it can contain and the length of time that information remains available. Second, there are different ways of testing an individual's memory performance, and some of these tests are more likely to detect existing memories than others. Third, the evidence supports the idea that there are at least two distinct memory systems, one that can be consciously summoned to complete a task at hand, the other a behind-the-scenes operator whose aid goes largely unnoticed. Fourth, circumstances such as illness, accident, and age can alter one's mnemonic abilities for the worse with sometimes devastating outcomes. Fifth, despite our vast reliance on and confidence in our recollections, memories can shift and become altered in both subtle and not-so-subtle ways. Finally, and perhaps most hopefully, there are a number of steps that can be taken to improve our memories.

Memory Systems

The explicit memory system is the system believed to aid people when they consciously attempt to rekindle memories of past events. For example, if you have ever tried to remember your first kiss, graduating from college, or a word from a list you studied 5 minutes ago—you have experienced your explicit memory system in action. Memory researchers have identified three types of memory that people can use when asked directly to recall information from the past. These types of memories are differentiated primarily on the basis of the amount of information that each "type" can hold (capacity), on how long a given memory can be expected to last (duration), and on the processes by which information is lost from the system (forgetting).

The type of memory with the shortest duration, but with a surprisingly large capacity, is known as the sensory memory system. Sensory memories are brief representations of the past. There are sensory memories for visual information (iconic memories), auditory information (echoic memories), as well as for the other perceptual senses such as touch and smell. The study of sensory memory is perhaps best exemplified by George Sperling's classic experiments using the partial report procedure. Sperling updated an older method called the whole report procedure that had previously been used to test participants' performance on visual sensory memory tasks. During a whole report trial, participants are presented with a very brief (on the order of 50 ms) presentation consisting of a matrix

of numbers or letters containing about 12 items. After the presentation, participants are simply asked to recall as many items from the matrix as possible. When confronted with this situation, participants are capable of producing, on average, three or four items.

The partial report procedure is similar but with one notable modification. In the partial report procedure, participants are again presented with a matrix of numbers or letters containing about 12 items for a very brief period of time. However, in the partial report procedure, instead of asking participants to recall all of the characters in the matrix, they are instead cued to remember only the top, middle, or bottom line of the matrix. In the standard experimental condition, participants do not know which line they will be asked to recall until after the matrix has been presented. In this case, participants are capable, on average, of recalling three or four items. However, because participants had no idea in advance which line of the matrix they would be asked to recall, they must have maintained a fleeting but fairly accurate memory representation for the entire 12-item matrix. Based on studies like Sperling's partial report procedure, it has been determined that sensory memories in all modalities are relatively brief, accurate representations of past events with a large capacity.

Although extremely brief memories might be beneficial for some purposes, in other cases it is necessary to maintain information over longer time intervals. For instance, if one were to look up a phone number, walk over to a phone and attempt to dial the number, a fleeting memory lasting less than a second would be of little use. Short-term memories provide us with the ability to maintain discrete amounts of information over time intervals on the order of minutes. Whereas the capacity of sensory memory is comparably large, research has indicated that individuals can hold only about seven plus or minus two "chunks" of information in short-term memory—a chunk being an arrangement of individual parts into a meaningful representation (for example, arranging the individual letters "D" "O" "G" to form the chunk "DOG").

Short-term memories hold a limited amount of information for up to several minutes, but how is that information lost? Originally, information was believed to be lost from short-term memory through a process of decay. In other words, after a stimulus is presented, one must constantly rehearse it or else it will gradually be lost as a result of the passage of time. Evidence for such an idea arose from experiments in which participants were asked to remember short sequences of letters such as "X" "Q" "B." After presenting the letters, participants were asked to count backwards by 3s starting at a number like "427" for varying time intervals ranging from 3 to 18 seconds. It was observed that the greater the amount of time that passed between study and test, the greater the amount of forgetting. However, these experiments traditionally employed multiple study/test trials with participants' memory performance displayed as an average across all of the trials. Subsequent examination of studies using paradigms such as this revealed that participants were almost always correct on the first trial—regardless of the delay interval. Thus, it was not the passage of time, per se, that led to forgetting but rather the interference compounded by multiple study/test trials. This basic result has been found to be true more generally as well.

Whereas the information in our sensory memory systems is available for, at most, only a few seconds and information in short-term memory is available for up to several minutes, information in long-term memory may be kept indefinitely. In addition, there are several types of long-term memory believed to be in operation, each comprising different types of information from the past. One type of information residing in long-term memory represents autobiographical information from an individual's past. These are known as episodic memories. Examples of episodic memories include instances such as remembering the time you attended a friend's wedding or the events that occurred on a vacation that you took last year. Another type of long-term memory represents knowledge about the past with no specific reference to a past autobiographical event. For example, it is unlikely that a person would remember the time that he or she learned his or her name or the day that the person learned that the capital of the United States was Washington, D.C. The type of long-term memory comprising general knowledge about the world is known as semantic memories. Finally, procedural memories are long-term memories typically consisting of learned motor skills. Remembering how to ride a bicycle or the motor movements necessary to swim in a pool are examples of this type of long-term memory.

Types of Memory Tests

Having described some of the different memory systems believed to be in operation, it is important to

note that there are several different types of memory tests—each different with respect to the likelihood that it will elicit a stored memory. These four types of tests are sometimes referred to as the four Rs, and they are recognition, reconstruction, relearning, and recall. A recognition test occurs when one attempts to choose a correct answer from among a number of distractors, such as on a multiple-choice exam. Although the manner in which information is encoded influences the type of test that is most likely to elicit a stored memory, in many circumstances, recognition tests are more sensitive than other types of memory tests. This means that in most circumstances, recognition tests are more likely to elicit stored memories than other types of tests.

Another type of memory test is a reconstruction test. A reconstruction test occurs when the test-taker studies some information, is provided with all of the studied information at the point of test, and is then asked to put that information back into its original order. For example, if you studied the words *cat*, *dog*, *mouse*, *pig*, and *giraffe* in that order, at the point of test, you might be given those same words back in a new random order (e.g., *pig*, *mouse*, *dog*, *giraffe*, and *cat*). In this instance, your task would be to put the words back into their original order.

Relearning tests, sometimes referred to as savings or savings in relearning, involve the fact that if a person learns something once, and forgets some of the information over time, it will typically take the person less time to learn it again. Work in this area of memory research was pioneered by Hermann Ebbinghaus. In one classic series of experiments that Ebbinghaus conducted with himself as a participant, he painstakingly memorized lists of nonsense syllables. After learning each list to a criterion point, typically the point at which the list could be repeated without mistakes, he would go back after a delay interval and study the list again until he could once again repeat the list without mistakes. What Ebbinghaus discovered is that with each relearning, it took him less and less time to reacquire the list, thereby demonstrating savings in relearning. A real-world example of this principle might be if you studied Spanish for several years in high school and were then required to complete more Spanish courses in college. It is probable that you would have forgotten some of the Spanish vocabulary between high school and college. However, according to the principle of savings in relearning, having memorized the information once would allow you to reacquire the forgotten Spanish vocabulary words more quickly than it took to learn them in the first place.

The fourth and final type of memory test to be discussed in this context is recall. Recall tests can be broken down into two basic categories—free recall and cued recall. A free recall test is one in which a person is asked to remember all of the studied information in the absence of any cues. For example, if a person read a short two-page story and was then presented with a piece of paper and a pen with the task being to reproduce the story in its entirety, that would be a free recall test. Free recall tests are often the least sensitive measure of memory performance. In other words, in most normal circumstances, they are the least likely to elicit stored memories. Cued recall tests, on the other hand, can be more sensitive than free recall tests. An example of a cued recall would be if one were asked questions about a story that one just read with the task being to recall specific events from the story. Although there are a number of circumstances in which cues do seem to aid recall performance, there are instances in which such cues fail to facilitate recall performance. One example of this is the part-set cueing effect. This rather counterintuitive effect occurs when a person studies a list of words and is then provided with several list items with the task being to remember the remaining items. Although one might expect list items in the form of cues to contribute to enhanced recall performance, they often have the opposite effect and render a person's memory for the remainder of the list items worse than it would have been in the absence of any cues at all.

Implicit Memory

Having discussed sensory memories, short-term memories, long-term memories, and several different ways of testing explicit memories, it is important to note that most memory researchers now agree that the explicit memory system is not the only system in operation. Again, explicit memory is the type of memory one employs if asked, for example, to purposefully retrieve information from a past event. Explicit memory tasks can be differentiated from implicit memory tasks in which memories from past events unconsciously affect memory performance in the present. One common way to reveal the implicit memory system in practice involves having participants complete priming experiments. There are two

main types of priming—repetition priming and association priming. However, despite the fact that both occur under different circumstances, they appear to operate similarly and have similar effects on memory performance. An example of a repetition priming task is the following: A participant in a memory study is tasked with looking at sequences of letters, in the form of words or nonwords, one at a time on a computer screen. The goal in this case is to determine as quickly as possible, for each amalgamation of letters presented, if it represents a word. If the word *doctor* was presented at some point during the letter presentation and then again at a later point, repetition priming occurs to the extent that *doctor* is recognized as a word more quickly the second time than the first time. Association priming, on the other hand, is what occurs when having seen *nurse* earlier in the list leads a participant to more quickly recognize *doctor* as a word than would otherwise be the case. Thus, the implicit memory system is unique in that it takes no conscious processing to occur.

Another distinction between implicit memory tasks and explicit memory tasks is differential task performance on measures of implicit and explicit memory. One classic anecdote reported by Edouard Claparede involves an account of a patient with anterograde amnesia (the inability to form new memories) who is unable to remember his doctor's name. One day, the patient's doctor hides a pin in his hand and pricks the amnesic patient. Later that day, despite the fact that the patient is unable to remember the event in which his hand was pricked, the patient is unwilling to shake the doctor's hand because of the realization that sometimes people hide pins in their hands. In this case, the patient's implicit memory system appears to be functioning relatively normally despite the complete and abject failure of the explicit system.

Memory Distortions

Given this description of anterograde amnesia, it should be noted that other forms of memory disorders can also occur. A slow and steady decrease in mnemonic abilities has been cited in some studies as a result of normal aging, but more abrupt and drastic changes can occur as a result of accident, disease, and other factors. In addition to the inability to hold on to new memories (the aforementioned anterograde amnesia), it is also possible for a person to lose memories from his or her past as the result of particular events (retrograde

amnesia). Although it is not uncommon for a person to lose a portion of his or her memory for events surrounding, for example, a traumatic head injury sustained in a car accident, the permanent loss of long-term memory is, fortunately, quite rare.

One critical and often misunderstood aspect of how memory functions is that memories are not like photographic snapshots of the past. Instead, when you conjure up an image of a robin, for example, the resultant memory is likely pieced together from a number of different past experiences. The unfortunate result of this process of reconstruction is that it is likely that some of the memories that we experience of past events are incorrect—in either subtle or more major ways. One clear example of this involves the case of flashbulb memories. Flashbulb memories are vivid accounts of a highly emotional or arousing event that engender a strong sense of confidence in the memory. For example, the *Challenger* shuttle explosion is one instance of an event that many people feel as if they can remember especially well. More recently, the events surrounding the terrorist attacks on September 11, 2001, seem to be painfully fresh and accurate in the minds of a number of people. But is the confidence in these types of memories justified? It turns out that despite people's confidence in their accuracy, memories for these types of significant and emotionally arousing events are surprisingly malleable and are prone to mistakes just like other memories.

One of the processes by which memory reconstruction appears to occur is through the use of schemas. A schema is an organized knowledge structure in long-term memory that can assist a person in reconstructing past events. However, although the use of schemas in memory reconstruction can be beneficial in terms of helping to aid accurate recall, they can also lead to problems and mistakes. For example in Sir Frederick Bartlett's classic experiments using the story "The War of the Ghosts," participants were asked to read a story containing a number of unfamiliar elements. When recounting the story after a period of delay, a number of participants supplanted some unfamiliar aspects of the story with more familiar ones ostensibly consistent with their schema for a story. In this manner, schemas may aid us in reconstructing events that might have occurred, but also lead us to make inaccurate inferences based on previous experiences.

Given that our memories for past events are reconstructed and are not snapshots of the past, is it possible to get a person to remember something that never happened at all? Unfortunately, the answer to this question

is that it is not only possible but also not especially difficult. Henry L. Roediger and Kathleen B. McDermott conducted one well-known laboratory demonstration of the ease with which false memories can be implanted in unsuspecting research participants. They had participants study lists comprising related words like *bed*, *rest*, *awake*, *snooze*, and *slumber*. After studying the lists, participants were asked to complete a free recall test for the studied information. Although it is relatively uncommon for participants to remember words on a free recall test that did not actually appear on the studied lists, Roediger and McDermott observed high levels of false recall for words like *sleep* that were highly related to the studied lists. Thus, despite the fact that this particular study involved word lists instead of other types of more commonly encountered stimuli, it was not difficult to get participants to remember events from the past that did not occur.

Improving Memory Performance

Despite the bad news about how memories can become warped and changed or even lost forever, there is some good news in that there are a number of readily available techniques that can allow one to improve upon his or her baseline level of memory performance. These mnemonic solutions are often closely related to the particular problem associated with memory that one is attempting to overcome. There is an often-cited trajectory that a memory takes from first experiencing an event to the ability to resurrect the memory for appropriate use at a later time. These stages of memory are (a) encoding or acquisition of the memory; (b) storage or maintenance in which memories are kept for later use; and (c) retrieval, or the point at which the memory is accessed. Problems with memory performance can occur at each of these three stages. For example, one problem that many people face with respect to memory performance is remembering the names of people whom they have recently met. In the specific case of memory for names, the most likely problem is that the name was not encoded properly in the first place. This could be a result of focusing on making a good first impression rather than focusing on the task of remembering the individual's name. If the locus of the problem is at the encoding stage, the solution likely involves encoding as well. Putting a greater focus on attempting to remember the name, even something as simple as repeating the name several times immediately after learning it, can vastly improve performance.

One problem often faced by students that is different from memorizing people's names is the task of attempting to understand, integrate, and store new information. In this case, the goal is to maintain a potentially large and varied amount of information in long-term memory. Although there are some technical mnemonic techniques that can be of use in these situations, the best solutions to these types of problems are perhaps both the most obvious and the most underappreciated. Technical mnemonics are memory tricks that one can use to memorize a given set of information more quickly and completely. For example, one method for learning vocabulary words from a foreign language involves the use of interactive visual images coupled with words from a language a person can already speak. Let's say that a person is attempting to learn a list of French vocabulary words, and one of the words on the list is the French word for *butter*, which is *beurre*. The French pronunciation of *beurre* sounds very much like the English pronunciation of the word *bear*. Therefore, if one were to form a mental image of a bear eating a stick of butter, they would be much more likely to remember the new word after a period of delay than if they had attempted to memorize the word using more traditional memory techniques, such as repetition.

However, despite the viability of using specific technical mnemonics in some situations, it is often difficult to apply them across a spectrum of varying mnemonic tasks. In spite of this rather severe limitation, there are at least two more general memory principles that can be of use, for example, to students attempting to memorize course material. One of these principles is that prior knowledge in a given field can make learning related material easier. For example, someone taking an introductory psychology course may find the course difficult because of the array of novel concepts being introduced, but he or she might find subsequent, ostensibly more difficult upper division courses easier than the introductory course because of his or her acquired knowledge in the field. John Bransford and Marcia Johnson conducted one well-known experiment demonstrating this concept. They had participants read steps consistent with washing clothes either with prior knowledge about the choice of topic or without such knowledge. They observed that memory performance for the information contained within the passage was much higher for participants provided with the opportunity to integrate the information contained in the passage into a prior knowledge framework than for those who were not.

In addition to prior knowledge increasing one's efficiency with respect to acquiring new information, the concept of massed practice versus spaced practice provides the tools for one of the best and most overlooked mnemonic strategies. Massed practice involves attempting to learn information over a relatively short time interval, whereas spaced practice involves breaking up study sessions into a number of discrete periods of learning. The results of studies examining memory performance under these two study conditions unequivocally support spaced practice over massed practice. In fact, in this instance, one could spend identical amounts of time going over the to-be-remembered information, and the mere introduction of intervals between study sessions is, by itself, enough to provide a major overall boost to memory performance.

Matthew Reysen

See also Long-Term Memory; Short-Term Memory

Further Readings

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology.* Cambridge, UK: Cambridge University Press.

Bransford, J. D., & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior, 11,* 717–726.

Claparede, E. (1951). *Reconnaissance et moiite.* In D. Rapaport (Ed.)*, Organization and pathology of thought.* New York: Columbia University Press.

Ebbinghaus, H. E. (1964). *Memory: A contribution to experimental psychology.* (H. A. Ruger, Trans.). New York: Dover.

Peterson, L. R., & Peterson, M. J. (1959). Short-term retention of individual items. *Journal of Experimental Psychology, 61,* 12–21.

Roediger, H. L., III, & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology, Learning, Memory, and Cognition, 21,* 803–814.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs, 74*(11).

MENTAL AGE

Mental age is a central concept in the study of intelligence measurement. Jerome Sattler defined mental age as "the degree of general mental ability possessed by the average child of a chronological age corresponding to the MA score" (p. 172). As an example, a child assessed with a mental age of 9 is viewed as having the general mental ability of an average 9-year-old child.

From the perspective of intelligence measurement, each individual has two ages: a chronological age that is the number of years that the individual has been alive, and a mental age that is the chronological age of persons for which the test performance of the individual is the average test performance.

The mental age and the chronological age of an individual need not be the same. For example, if the mental age of an individual is greater than the chronological age of the individual, then one can infer that the individual has above-average intelligence or higher mental ability.

The mental age and the chronological age for an individual are used to determine the ratio IQ (intelligence quotient) of the individual. To compute the ratio IQ, one divides the mental age of an individual by the chronological age of the same individual and then multiplies that ratio by 100. For example, if a child has a mental age of 12 and a chronological age of 10, then the ratio IQ for that 10-year-old child is 120 (i.e., $12/10 \times 100 = 120$).

The ratio IQ was the measure of intelligence used in the 1916 and 1937 versions of the Stanford–Binet Intelligence Scale and on other tests of mental ability. The deviation IQ replaced the ratio IQ as the measure of intelligence used in subsequent measures of intelligence. The deviation IQ reflects the location of the test performance of an individual in a distribution of the test performances of other persons with the same chronological age as the individual, with the mean deviation IQ being typically equal to 100. For example, if an individual has a test performance that is less than the mean test performance for same-age peers, then the individual will have a deviation IQ less than 100. Neither the dated measure of ratio IQ nor the more contemporary measure of deviation IQ consistently provides concrete information as to the reasoning skills of individuals.

The mental age score may also be termed the *age-equivalent score* according to Sattler. The mental age score for an individual provides information as to what age group is most closely associated with the individual from the perspective of mental ability. As an example, a 12-year-old child with a mental age score of 14 indicates that the 12-year-old child has

a mental ability more typical of 14-year-old children than of 12-year-old children.

Sattler noted that mental age scores have certain limitations. First, differences in mental age do not reflect the same differences in mental ability across the age spectrum. For example, the difference in mental ability between a mental age score of 5 and a mental age score of 2 tends to be greater than the difference in mental ability between a mental age score of 15 and a mental age score of 12. Second, the same mental age may reflect different capabilities for different individuals. For example, two children both with the mental age score of 12 may have answered different test items correctly.

Louis Thurstone was highly critical of the mental age concept. Thurstone argued that "the mental age concept is a failure in that it leads to ambiguities and inconsistencies" (p. 268). To Thurstone, mental age may be defined in two different ways. The mental age of an individual may be defined as the chronological age for which the test performance of the individual is average. The mental age of an individual may also be defined as the average chronological age of people who recorded the same test performance as the individual. To Thurstone, these two definitions do not engender the same numerical scores. In addition, if one accepts the first definition, one faces the problem that there may be many chronological ages for which a test performance is average. For example, a 16-year-old adolescent who provides a typical test performance for 16-year-old adolescents could be viewed as having a mental age of any score from an adolescent mental age score of 16 to an adult mental age score of 40. The average mental test performances of older adolescents and adults tend to be very similar.

Thurstone did not support the continued use of mental age or IQ as a measure of intelligence. However, he did support the use of percentiles for same-age peers in designating personal mental abilities. For example, if the test performance of a 12-year-old child receives a score that is equal to the score of the median test performance among 12-year-olds, then that 12-year-old child may be viewed as receiving a percentile of 50 (i.e., the test performance of the 12-year-old child is equal to or greater than 50% of the test performances of all of the 12-year-old children who were tested).

Despite the trenchant criticism of the mental age concept by Thurstone and the recognized limitations of mental age scores, noted commentators on intelligence such as Sattler and Lloyd Humphreys extolled the merits of the mental age score as an informative measure of mental ability. Both Sattler and Humphreys contended that the mental age score provides useful information about the mental capabilities of an individual. The mental age score provides information about the size and the level of maturity of the mental capabilities of an individual. The IQ score, whether the ratio IQ or the deviation IQ, provides no such information.

Both Sattler and Humphreys contended that mental age will likely continue to be a popular and useful measure of mental ability. However, the suggestion by Thurstone that percentiles among same-age peers be used to index mental abilities continues to be worthy of further consideration. Only time will tell whether the percentile or some other index will replace mental age as a popular index of mental ability.

Wiliam M. Bart

See also Intelligence Quotient (IQ); Intelligence Tests; Measurement

Further Readings

Humphreys, L. G. (1985). General intelligence: An integration of factor, test, and simplex theory. In B. B. Wolman (Ed.), *Handbook of intelligence: Theories, measurements, and applications* (pp. 201–224). New York: Wiley.

Sattler, J. M. (2001). *Assessment of children: Cognitive applications* (4th ed.). San Diego, CA: Jerome M. Sattler, Publisher, Inc.

Thurstone, L. L. (1926). The mental age concept. *Psychological Review*, 33, 268–278.

MENTAL HEALTH CARE IN SCHOOLS

The primary function of schools is to educate students, but they are also responsible for helping to nurture positive social, behavioral, and psychological development in children and adolescents. Many youth attending today's schools face a number of environmental, social, and economic challenges that create stress and often cause significant distraction from the primary task of learning. Educators and health professionals now understand that unresolved physical, mental, or psychosocial problems become barriers to effective learning and that school staff must facilitate access to mental health interventions if they hope to

see the majority of their students succeed. School-based and school-connected mental health services offer students the possibility of taking full advantage of the educational opportunities available to them.

Current statistics indicate that the prevalence of mental health problems experienced by youth has reached crisis proportions. Plans to improve children's mental health must acknowledge the important role that schools play in ameliorating the impact that stress has on children and adolescents. School mental health services address the barriers to care related to cost, availability, and the accessibility of a continuum of services, challenges inherent in community-based mental health organizations. Schools have provided support services and programs through school counselors, school social workers, school psychologists, and school nurses for many decades, yet the number of available professionals employed by schools has not been adequate to meet the level of need demonstrated by the general education population. Promising and effective solutions include the development of school-community partnerships and the implementation of comprehensive, school-based mental health services. These approaches address the growing need for mental health care and increase the likelihood that youth can benefit from the learning opportunities available to them.

Estimates of Mental Health Needs Among Children and Youth

Children's mental health problems have reached alarming proportions, and many would consider present statistics to signify a national crisis. The levels of violence, maltreatment, and poverty plaguing some homes and communities have a negative impact on the vulnerable and maturing systems of children and adolescents. One in five youth in the United States today experience mental health symptoms that cause some level of impairment, and 5%–10% of the population actually experiences significant dysfunction in one or more areas of daily living caused by a psychiatric problem.

A number of mental disorders with onset in childhood and adolescence require significant intervention by mental health professionals in order to reduce the risks and the consequences of these conditions. Prevalence rates for disorders of childhood and adolescence suggest that 3%–5% of school-aged children have attention deficit hyperactivity disorder, 5% of adolescents have major depression, and 13% of

children from 9–17 years of age have some form of anxiety disorder. Disruptive behavior disorders, mood disorders, anxiety disorders, eating disorders, and substance use disorders represent mental illnesses that may afflict an individual during his or her school-aged years, but the incidence of these pathologies, compared to other psychosocial problems, remains relatively low. The most common mental health problems reported among students include social, interpersonal, and family problems. In reality, most young people do not have a diagnosable disorder but do exhibit a subset of symptoms that cause enough suffering to warrant clinical attention and intervention.

Data indicate that the majority of youth with mental health concerns (approximately 80%) do not receive the services they need. Public mental health agencies have not effectively reached the majority of children and adolescents needing clinical care. Few community mental health systems have successfully addressed the barriers that often prevent access to services, such as lack of transportation, the cost and availability of mental health services, the lack of comprehensive coverage provided through insurance plans, the diminishing streams of public funds available, and the stigma still associated with being the recipient of mental health care. Furthermore, the public system of mental health care for children and youth is fashioned after the medical model and relies on assigning a psychiatric diagnosis to children in order to access appropriate interventions. Such stipulations, coupled with a characteristically high no-show rate, have exaggerated the gap between mental health needs and services.

Access and Utilization of Mental Health Services

Studies demonstrate that bringing mental health services to schools improves access to care and decreases the stigma associated with receiving needed support. The Surgeon General's Report of 1999 on children's mental health and the President's New Freedom Commission Report released in 2003 emphasize the public health field's involvement in addressing the growing need for supportive interventions and point to a specific role that schools can play in improving the identification, prevention, early intervention, and treatment of mental health problems. Schools number more than 100,000 across the United

States, and they are logical places to provide mental health interventions for many children and youth. They have been a site for some level of mental health care for numerous decades. They offer a unique and valuable opportunity to effectively address the emotional and behavioral symptoms students exhibit, given that youth spend most of their days in schools, form relationships with caring school staff, and can be assessed in multiple settings across various points in time. Mental health care conducted in schools is usually provided at no cost to families, and schools are especially sensible sites for implementing mental health services for youth and families who might be disengaged from traditional community-based mental health service systems.

For many years, special education programs have been the sole avenue for the delivery of mental health services in public schools for students deemed "emotionally disturbed." The Individuals with Disabilities Education Act (IDEA) was enacted as federal legislation to require states to provide any supports needed to guarantee an equal educational opportunity for students with documented disabilities. Although estimates vary state by state, approximately 9% of children ages 6 to 21 years of age nationally struggle with mental health problems that interfere with their academic performance and entitle them to services outlined under IDEA. All public schools are required by law to provide or facilitate treatment or counseling for students with mental health conditions if they have an Individualized Education Program (IEP) for an emotional disturbance, but there are few mandates for schools as they relate to students in the general student population who have mental health needs that do not directly impair their academic performance or that are subclinical in nature. Although not required by federal or state law, many schools recognize the need to make available a variety of mental health services and programs, including prevention, early intervention, and treatment services, to all students enrolled in a school if it is thought that they might benefit from them and improve educational outcomes.

More specificity is emerging regarding how schools could address the mental health crisis affecting children and youth. President George W. Bush established the New Freedom Commission on Mental Health in 2002 to examine the gaps in the public mental health system and to provide recommendations on how to improve the quality of available mental health services throughout the country. The Commission's findings concluded that the national mental health system required a significant transformation and that schools are in an important position to deliver or facilitate services for children and adolescents exhibiting mental health concerns. The Commission advocated for an increase in mental health programs in schools that would make empirically supported screening approaches, prevention programs, and treatment services available to all students in need, either through on-site services or referrals to community providers. They recommended that the country improve and expand school mental health programs as a key service delivery model to address the nation's public health concerns.

School-Hired Mental Health Professionals

Research now confirms that public school personnel are the major providers of mental health services for school-aged children. School support staff, including school counselors, social workers, teachers, and administrators, form a group that meets regularly to assist students in regular education who are having difficulty in the school environment. These early intervention teams (also referred to as student support or child study teams) strive to intervene as early as possible when there is concern about a student's success and engage in a problem-solving process to identify and implement viable solutions.

School Counselor

School counselors are the most common type of school-hired mental health professional, with more than 100,000 estimated to be working in public schools across the United States. School counselors generally help all students enrolled in a school in the areas of academic achievement and personal/social development, and they provide mental health education, early identification, and intervention services for students in need and their families. As a result of their advanced training (a master's in education or counseling) and the completion of the respective state licensure requirement, school counselors are qualified to provide direct mental health services such as individual and group counseling, behavioral interventions such as skill development and goal setting, and crisis intervention services to identified students. Although the American School Counselor Association cautions school administrators about the use of

school counselors for inappropriate program tasks, they are too often pulled to perform other organizational or management responsibilities such as conducting academic, career, or vocational counseling; monitoring attendance or disciplinary problems; or maintaining student schedules or academic records, thus limiting their time for direct mental health interventions.

School Social Worker

Most school social workers have at least a master's degree in social work and operate as a link between the home, the school, and the community for all students enrolled in a school. Estimates differ, but there are approximately 15,000 social workers involved in public school systems across the country. School social workers are generally trained to conduct a number of mental health services for students (i.e., individual and group counseling, psychosocial assessments); families (i.e., consultation and case management); school professionals (i.e., training and school-wide prevention); and the community (i.e., referrals and collaboration). In reality, significant pressure felt by school administrators to comply with federal mandates such as IDEA and Section 504, the continued growth in the number of students entering the special education system, and the limited number of qualified providers available to conduct the services outlined within an IEP often lead to the exclusive use of school social workers for students with identified disabilities.

School Psychologist

There is greater variability state to state with regard to educational and training requirements for school psychologists. School psychologists perform a number of therapeutic services and are trained to assess and evaluate a wide variety of behavioral problems, skills deficits, emotional conditions, and cognitive abilities. As a result, they play a central role in determining who is eligible for special education; administering IQ, personality, and achievement tests; and conducting therapy with students displaying significant mental health problems. Given the small number of school psychologists that may be found in any particular school system, the higher salary they generally draw, and their specialized skill sets, school psychologists often serve more than one school during the year. School districts

usually cannot afford to use school psychologists for services that are not reimbursable (i.e., prevention) or mandated (i.e., not part of an IEP).

School Nurse

There are approximately 45,000 school nurses employed to work in public schools. Although not often associated with mental health interventions, school nurses play a role in addressing the mental health needs of children and adolescents in many schools. Surveys indicate that more than one third of the school nurse's time is devoted to mental health service provision. With appropriate training, the school nurse can be involved in prevention, early intervention, and even treatment activities that support youth with mental health concerns. It is not uncommon for youth to mask their mental health concerns by complaining of physical symptoms such as stomachaches or headaches. Under such circumstances, the school nurse may be the first helping professional a student will see and therefore can screen and identify those who may actually be experiencing psychological distress. School nurses play a role in the delivery of early intervention services, particularly around skill development and the provision of support for psychosocial issues. They are also important partners in the delivery of mental health treatment for children who are required to take psychotropic medications for serious mental health conditions.

School-Community Partnerships

Public schools hire a fair number of professionals to focus on the mental health needs of students, yet few schools have sufficient staff to address the plethora of emotional and behavioral issues exhibited among the student population. Because school administrators recognize that they must address the frequency, intensity, and diversity of mental health needs of their students, formal arrangements between schools and community mental health agencies, health departments, or local hospitals are becoming more common across the United States. Inviting qualified, community-based providers to work in schools increases the clinical capacity of schools to provide mental health services and supports to all students in need and extends the public mental health system of care to include school communities. Building school-community collaborations broadens the continuum of care; creates a stronger

and more reliable network of support; increases the utilization of services; and improves access to interventions, especially among adolescents, a subpopulation that rarely uses community-based health services.

School-based health centers (SBHCs) represent a partnership between a school, community, and health provider; are located on or near school grounds; and often deliver an array of physical, mental, and dental health preventive and treatment services to children and adolescents. There are currently 1,725 SBHCs operating in or linked to schools. About 10% of schools have an SBHC that offers mental health and social supports to enrolled students.

Although these collaborative agreements improve access to mental health care, they are not without their limitations. A greater number of mental health contacts do not signify greater quality of care, suggesting that attention to continuous quality improvement activities remains critical. Community mental health organizations are often Medicaid or insurance driven, meaning that clients must be diagnosed in order for payment sources to be accessed. This may lead to a heavier focus on treatment services and less of an emphasis on prevention or early intervention services. Very few community mental health providers are familiar with the nuances of working in schools, do not recognize or respect the norms of school cultures, and are unaware of the goals and objectives that drive schools. Furthermore, training and experience among community mental health providers is usually traditional, leading individuals to rely on knowledge gained from office-based practices as opposed to assuming flexible approaches to the provision of services that are required for school-based interventions to be successful. In spite of these challenges, there are a number of notable school-based mental health programs that have effectively trained community mental health providers to work collaboratively in school settings and have demonstrated positive outcomes among students, staff, and families.

Continuum of Mental Health Services

A number of mental health interventions can be offered in schools, and a variety of professionals are qualified to deliver psychological interventions, counseling, behavioral supports, and social services. School-based mental health programs have become as diverse as the students they serve and the schools within which they operate. Professionals in the field have identified the

key components that define comprehensive, school-based mental health programs. These programs usually include mental health promotion and prevention activities; screening, identification, and early intervention; treatment and referral services; and crisis intervention.

Mental Health Promotion and Prevention

Childhood and adolescence are important times to teach healthy forms of coping with normal or abnormal stressors. Mental health promotion and prevention activities both use approaches that engage entire populations to impede the development of mental health problems, yet there are some subtle differences between these two concepts. Mental health promotion aims to identify and support the determinants of health—both the behaviors that allow individuals to maintain healthy lifestyles and the environments that support optimal health. Mental health promotion activities used in schools would be those that strengthen individual competencies or increase protective factors and foster nurturing environments, such as interventions that promote self-esteem, improve school climate, encourage healthy relationships, and build networks of care.

Prevention interventions, specifically those called universal interventions, identify factors that put entire populations at greater risk for the development of mental illness or psychological problems and implement strategies to reduce those risks. These interventions prevent the onset of dysfunctional patterns of behavior that often lead to negative outcomes. Selective prevention programs, on the other hand, address the needs of those students who are at the highest risk of problematic behavior. A large number of school-based programs have been developed to prevent a variety of problematic outcomes exhibited among an increasing number of youth, such as substance use, teen pregnancy, and interpersonal violence. Teachers, administrators, other school staff, and mental health professionals have been involved in the implementation of prevention and promotion activities delivered in schools.

Screening, Identification, and Early Intervention

Youth are in schools for the majority of their day, making schools an excellent place to identify those who are beginning to exhibit symptoms of mental health problems. School staff can be trained to implement school-wide, classroom-based, or individually

administered screening tools to identify early warning signs of emotional or behavioral disturbance. The President's New Freedom Commission concluded that an effective mental health system would include mechanisms for early detection of mental health problems in children and suggested that interventions should be available in low-stigma settings, such as schools. Early intervention services are developed to intervene early after a student is identified as demonstrating a behavioral or psychological symptom. Intervening after symptoms are first apparent can change the course of some mental health conditions and reduce the long-term consequences associated with mental illness. Examples of early intervention strategies for school-aged youth include anger management groups, social skills groups, and teacher consultation regarding mental health issues. Screening and the delivery of early intervention activities requires some training in mental health or counseling; therefore, school-hired or community mental health professionals are usually the most appropriate individuals to lead such activities.

Treatment and Referral Services

Some students develop more significant mental health problems that require coordinated care and greater clinical expertise. The symptoms students exhibit may or may not meet criteria for a mental health diagnosis, but treatment services offered in schools can help prevent the worsening of symptoms and the need for more intensive services. Symptoms related to depression, trauma, and grief or loss usually indicate a need for greater clinical intervention. Modalities of treatment for more serious mental health problems include individual, family, or group therapy; substance abuse counseling; and psychopharmacological treatment. Mental health providers, either school-hired or community-hired, should be trained in child and adolescent treatment approaches; hold a license to practice therapy, counseling, or social work; or be well supervised by a licensed provider. If school-based treatment is not appropriate, school mental health professionals can refer students and their families to community-based or hospital-based mental health programs offered outside of school.

Crisis Intervention

During emergency situations, as a result of an unpredictable acute event or in reaction to high levels of stress, students sometimes display severe reactions that require well-coordinated responses from school staff. Interpersonal crisis situations that occur in schools include threats of suicide or homicide, physical assaults, and severe behavioral acting out. School-based crisis intervention approaches involve specific education, ongoing training, and established protocols to be most effective. Similar to the qualifications necessary for treatment providers, those leading responses to mental health crises should be trained in child or adolescent mental health and have obtained advanced training in crisis intervention methods.

Conclusion

Current statistics indicate that the prevalence of mental health problems experienced by youth has reached crisis proportions. Plans to improve children's mental health must acknowledge the important role that schools play in ameliorating the impact that stress has on children and adolescents. School mental health services address the barriers to care related to cost, availability, and the accessibility of a continuum of services—challenges inherent in community-based mental health organizations. Schools have provided support services and programs through school counselors, school social workers, school psychologists, and school nurses for many decades, yet the number of available professionals employed by schools has not been adequate to meet the level of need demonstrated by the general education population. Promising and effective solutions include the development of school-community partnerships and the implementation of comprehensive, school-based mental health services. These approaches address the growing need for mental health care and increase the likelihood that youth can benefit from the learning opportunities available to them.

Olga Acosta Price

See also *Diagnostic and Statistical Manual of Mental Disorders*; Individualized Education Program

Further Readings

Acosta, O. M., Tashman, N. A., Prodente, C., & Proescher, E. (2002). Establishing successful school mental health programs: Guidelines and recommendations. In H. Ghuman, M. Weist, & R. Sarles (Eds.), *Providing mental health services to youth where they are: School and community-based approaches*. New York: Brunner-Routledge.

Adelman, H. S., & Taylor, L. (2006). Mental health in schools and public health. *Public Health Reports*, *121*, 294–298.

Brener, N. D., Martindale, J., & Weist, M. D. (2001). Mental health and social services: Results from the School Health Policies and Programs Study 2000. *Journal of School Health*, *71*(7), 305–312.

Foster, S., Rollefson, M., Doksum, T., Noonan, D., Robinson, G., & Teich, J. (2005). *School mental health services in the United States: 2002–2003* [DHHS Pub. No. (SMA) 05–4068]. Rockville, MD: Center for Mental Health Services, Substance Abuse and Mental Health Services Administration.

Kutash, K., Duchnowski, A. J., & Lynn, N. (2006). *School-based mental health: An empirical guide for decision-makers.* Tampa: University of South Florida, The Louis de la Parte Florida Mental Health Institute, Department of Child and Family Studies, Research and Training Center for Children's Mental Health.

New Freedom Commission on Mental Health. (2003). *Achieving the promise: Transforming mental health care in America: Final report* (DHHS Pub. No. SMH-03–3832). Rockville, MD: Author.

Rones, M., & Hoagwood, K. (2000). School-based mental health services: A research review. *Clinical Child and Family Psychology Review*, *3*(4), 223–241.

U.S. Department of Health and Human Services. (1999). *Mental health: A report of the Surgeon General.* Rockville, MD: Author.

Weist, M. D. (1997). Expanded school mental health services: A national movement in progress. In T. H. Ollendick & R. Prinz (Eds.), *Advances in clinical child psychology.* New York: Plenum.

MENTAL RETARDATION

Mental retardation is a disability characterized by significant limitations in both intellectual functioning and adaptive behavior. The disability originates in the developmental period before the age of 18. Individuals with mental retardation are a heterogeneous group. There is no single cause of the disability, and areas of strengths and weaknesses can differ widely from individual to individual. Most individuals with this disability have IQs in the higher ranges of mental retardation, and their disability may be initially noticed primarily on academic tasks.

Between 1% and 3% of the total population in the United States has mental retardation. It is difficult to precisely calculate the number because individuals are not routinely screened, preferred terms for the disability may vary among jurisdictions, and states sometimes use different eligibility cutoffs for different purposes (e.g., eligibility for special education purposes may be different from eligibility for subsidized housing). Additionally, some individuals who might meet the criteria for a diagnosis of mental retardation purposefully avoid the diagnosis and stigma by attempting to appear mentally typical in their community and workplaces. Consequently, the number of adults officially reported to have mental retardation is likely lower than the actual number.

It is understandable that many individuals would not want to be identified as having mental retardation because being labeled with this disability exposes people to the discriminatory attitudes of society. Sadly, people perceived to have an intellectual disability are often socially constructed in ways that stigmatize them and limit their opportunities for full participation in schools, communities, and workplaces. There is evidence that societal attitudes toward persons with intellectual disability are changing, but many prejudices and stereotypes associated with this disability remain.

History and Terminology

Societies have always recognized differences in abilities between individuals. Throughout the ages, people have attempted to describe and name these differences. The earliest description of mental retardation was found written on ancient papyri in Thebes (1500 B.C.), but it is likely that people have been attempting to understand and name observed differences in human capabilities since the origins of civilization.

People's views toward and treatment of individuals seen as differing significantly from the norm or as having disabilities have fluctuated across the years depending on particular circumstances within a given society. Religious beliefs, cultural values, and economic conditions are some of the factors that influence a society's view of disability. There is evidence, for example, that some early societies, such as the Egyptians, highly valued children and spent much effort on nurturing and caring for them. This included children with disabilities. Other societies, such as the Greeks and Romans, respected physical prowess and practiced infanticide when children were born with a clearly distinguishable disability.

Perceptions of individuals with disabilities in more recent times have also shown great variation. At certain points in time, such as in the early 19th century,

society in general promoted humane treatment for persons with disabilities. During that time, many professionals in Europe and the United States began to develop educational programs for individuals with mental retardation and other disabilities. When it became clear that treatment could not "cure" disability, and the economic status of the United States changed during the latter part of that same century because of the Civil War, attitudes toward persons with disabilities shifted. They became seen as burdens to society, incapable of benefiting from educational programs or contributing to their communities. Large-scale institutions where people with disabilities were segregated and that provided only minimal care became the norm at this point rather than the progressive programs seen earlier. The trend toward institutionalization continued throughout the early part of the 20th century.

In the mid- and latter part of the 20th century, legislation and civil rights litigation by individuals with disabilities and their families created major shifts in service provision and in the ways in which individuals with disabilities are portrayed and viewed by society. This resulted in improved access to higher-quality education and supports for persons with mental retardation within inclusive rather than segregated settings, and increased opportunities to be fully participating, valued members of their families, schools, and communities.

The principle of normalization, introduced from Scandinavia by Bengt Nirje in the late 1960s, also had an enormous influence on the field. This principle stated that persons with mental retardation have the right to the same everyday social, economic, and personal life experiences as their typical age peers, within the same settings used by other community members. The principle also emphasized the right of individuals to have their choices and preferences considered in decision making. Application of the normalization principle in service systems formed the basis of the movement to bring people with mental retardation out of institutional settings and assist them to become active participants in their communities.

The terms applied to individuals who exhibit characteristics associated with mental retardation have also changed throughout the years. *Feebleminded, simpleton, fool, idiot, imbecile*, and *moron* are some of the names that have been used in the past to refer to persons perceived as having a cognitive disability. Some of these names were originally scientific terms used by professionals in the disability field (e.g., moron), but

over time, the terms came to be viewed as derogatory and were replaced with other, less offensive terms.

Today, there continues to be variation in terminology to describe mental retardation. In Great Britain, the term most widely used is *learning disability;* in other parts of the world, the term *intellectual disability* is preferred. *Mental retardation* is currently the most widely used term within the United States, although *mental disability, intellectual disability*, and *cognitive disability* are gaining use. Like names from earlier eras, the term *mental retardation* has come to be seen as stigmatizing, and individuals with disabilities and their families are increasingly vocal about their objection to its use. Although no consensus has been reached, it is likely that the term *mental retardation* will be replaced with the synonym *intellectual disability.*

Definition

Just as the terminology used to refer to mental retardation has changed across time, so have the criteria used to define this disability changed as theories of intelligence, techniques of measurement, and the conceptualization of mental retardation have evolved. Early definitions of mental retardation focused on what was seen as the incurability of the disability. These definitions placed more emphasis on measurement of intelligence (i.e., IQ score) in determining if someone had an intellectual disability than on other aspects. Over time, there has been increasing recognition of the importance of measuring an individual's adaptive behavior skills and current functioning in determining a diagnosis. Currently, there is emphasis on an assessment process that considers intellectual ability, adaptive behavior, age of manifestation, and functioning.

By 1977, most definitions of mental retardation acknowledged the importance of clinical judgment. This is a particular type of judgment based on data and on a professional's extensive experience and expertise. Current definitions of mental retardation specify that an individual's score on an appropriate measure of intelligence fall at or below two standard deviations below the mean (i.e., at or below 70), considering standard error of measurement and the instrument's other properties. Use of clinical judgment is especially important for determining a diagnosis when an individual scores at the margins.

By the early 1990s, there was recognition that an individual's current level of functioning is affected by interaction between an individual's capabilities and

other factors, such as his or her participation, interactions, and social roles; health; and the context or environment in which he or she lives. Based largely on this understanding, the American Association on Mental Retardation (AAMR; now named the American Association on Intellectual and Developmental Disabilities, or AAIDD) made a major change in 1992 when it issued a new definition that included a classification system that, instead of subdividing the individuals into categories based on IQ score ranges, subdivided (classified) based on the intensity of a person's needs for supports. Development of the 1992 definition was also the first time that individuals with mental retardation were included in the discussions around the definition. Their inclusion as a part of the process of creating a new definition of intellectual disability is indicative of the growing strength of the self-advocacy movement among persons with disabilities. The AAMR made additional changes to the definition in 2002, primarily adjusting the manner in which adaptive behavior was conceptualized and assessed.

Today, discussions continue on how to precisely define and measure mental retardation. In general, however, all current definitions of mental retardation specify that an individual must demonstrate significant impairment in intellectual functioning and in adaptive behavior and that the disability must begin in the developmental period if he or she is to be considered as having mental retardation. The most widely used definition is the 2002 AAMR definition written by Luckasson et al.:

> Mental retardation is a disability characterized by significant limitations both in intellectual functioning and in adaptive behavior as expressed in conceptual, social, and practical adaptive skills. This disability originates before 18. (p. 1)

AAMR also stipulates that certain assumptions are essential to the application of the definition. An individual's functioning must be measured against typical community-based functioning of his or her age peers. The evaluator must consider an individual's skills within the context of his or her culture and language, and also take into consideration any behavioral, communication, sensory, and motor differences that might affect how the individual currently functions. It must also be recognized that each individual has strengths as well as areas of weakness, and that given appropriate supports, an individual's ability to participate over time in his or her home, school, community, and workplace generally will improve.

Two other widely used definitions are those found in the American Psychiatric Association's *Diagnostic and Statistical Manual of Mental Disorders* (Fourth Edition, Text Revision) *(DSM–IV–TR)*, and in the World Health Organization's *International Statistical Classification of Diseases and Related Health Problems* (10th ed.) *(ICD-10)*. The *DSM–IV–TR* definition is closely modeled on the AAMR 1992 definition. This definition specifies that an individual with mental retardation functions significantly below average in intelligence and adaptive behavior and that the age of onset is prior to age 18. This definition, as well as the *ICD-10* definition, differ from the AAMR definition by retaining a classification system based on IQ score ranges.

All three definitions reflect that mental retardation is a complex construct that must be interpreted within an ecological perspective. Thinking of mental retardation as a disease or condition is too simplistic and ignores the mediating effects of supports. Human functioning does not remain static but changes across time because of a number of interrelated factors, including the demands of environments; the types of supports and resources available to provide assistance; health status; and the opportunities available for participation and interaction within home, community, and work settings.

Diagnosis, Assessment, and Classification

Assessment for the purpose of diagnosis should result in benefit for the individual. This is especially true when considering assessment that may result in application of a potentially stigmatizing label, such as mental retardation. As discussed above, society's construction of mental retardation may have damaging effects on individuals carrying this label. Failing to detect cognitive disability when it is actually present, however, might cause an individual to lose access to critical resources and supports that would enhance life functioning and ensure fair treatment by society. On the other hand, applying an incorrect label of mental retardation when cognitive disability is not present may cause the individual to be stigmatized and to lose access to critical social, educational, and vocational opportunities.

Diagnosis

To be given a diagnosis of mental retardation, an individual must score at least two standard deviations

below the mean on an individually administered, standardized test of intelligence and score two standard deviations below the mean on a standardized assessment of adaptive behavior. In addition, the disability must have originated during the developmental period.

The purpose of diagnosis should be to determine eligibility for benefits, services, and supports. For example, children who have mental retardation and need special education qualify for specialized educational supports and services under the Individuals with Disabilities Education Act 2004 (IDEA). Once it is established that a child has intellectual disability, an individualized education plan (IEP) can be developed for the child to ensure that he or she is provided with the necessary educational interventions and related services to access and make progress in the general curriculum. Another purpose for diagnosis is decision making. In the criminal justice system, for instance, it is critical to determine if a defendant has mental retardation in order to ensure that the law is correctly applied in a given situation (e.g., in cases involving consideration of the death penalty).

Assessment

Valid assessment to determine if someone has an intellectual disability must include several factors. First, the examiner should have specialized training in administration of the assessment instruments required for establishing a diagnosis of mental retardation. The examiner should also have direct, extensive experience with individuals with mental retardation so that he or she can apply appropriate clinical judgment to the assessment process as warranted.

A second critical factor for valid assessment is that appropriate assessment instruments be selected and administered. To appropriately assess intelligence, the instrument selected must be standardized, individually administered, and based on a normative sample that included sufficient numbers of individuals across the age span who represent diverse groups of people (e.g., varied ethnic populations, abilities, rural and urban populations). Cultural and linguistic diversity are foundational considerations and must be considered in planning and implementing all aspects of assessment.

It must be kept in mind, however, that standardized assessments of intelligence were not designed to adequately evaluate populations of individuals who score significantly above or below the mean. Because, by definition, this includes individuals with intellectual disability, their scores on such instruments must be interpreted cautiously. It is also essential to remember that all assessment instruments are affected by errors in measurement. The standard error of measurement for most intelligence assessments is ± 3–5 points. For individuals scoring at the margins (i.e., at or close to 70), this consideration is especially important, and the examiner must use additional information (e.g., measures of adaptive behavior) and clinical judgment when making a diagnosis determination.

An additional consideration is that many individuals who may have mental retardation also have accompanying physical and sensory challenges. For example, they may not be able to complete test items requiring manipulation of testing materials or may need accommodations to access test stimuli (e.g., enlarged print). This factor underscores the necessity of having an experienced evaluator administer and interpret assessments that might lead to a diagnosis of mental retardation.

The second criterion for diagnosis of mental retardation is significant limitations in adaptive behavior skills. Adaptive behavior refers to the skills needed to be successful in one's everyday life. It includes multiple components, such as daily living, social, and conceptual skills. Valid assessment requires administration of a standardized measure of adaptive behavior. One of the most commonly used instruments is the Vineland Adaptive Behavior Scales. To meet the criteria for diagnosis of mental retardation, an individual must score at least two standard deviations below the mean in one of three broad skill domains of adaptive behavior (i.e., conceptual, social, or practical adaptive skills) or have an overall total score that is a minimum of two standard deviations below the mean. The instrument used for assessment must have been normed on diverse individuals across the age span and on persons with and without mental retardation.

Like intelligence, adaptive behavior is a complex construct, and it is difficult to develop an instrument that adequately assesses all adaptive behavior skills. Just as with assessment of intelligence, it is essential that the examiner have extensive experience both with appropriate adaptive behavior instruments and with individuals with mental retardation so that suitable clinical judgment can be applied as needed. When interpreting scores on adaptive behavior instruments, the examiner must also consider accompanying sensory or motor impairments and cultural, linguistic,

and socioeconomic factors that might affect the individual's performance of adaptive behavior skills.

Classification

One of the primary reasons to classify or subdivide individuals with mental retardation is to assist with planning services and supports. Various systems of classification have been used depending on the purpose for classifying. The mild/moderate/severe/profound system based on IQ ranges is still used in the *DSM–IV–TR* and the *ICD-10* definitions.

In 1992, the AAMR committee, working on a new definition of mental retardation, eliminated these IQ-based categories and instead recommended that individuals be classified by their need for supports. Supports encompass a range of activities, resources, and strategies that assist a person in any and all areas of his or her life. Supports can be provided by self (e.g., using a self-management checklist to complete job duties) or by other people (e.g., a teacher providing literacy instruction). They can include adaptive equipment or technology that allow access to information or enhance participation in an activity. Supports may also be some type of service; for example, case coordination to assist with performing routine activities such as paying bills or buying groceries.

The underlying concept on which this supports-based model is built is that when individuals with mental retardation (or, indeed, anyone) are provided with the supports they need, their functioning will likely improve. This concept represents a major shift from earlier thinking that viewed individuals with mental retardation as incapable of making substantive positive changes. The new model recognizes that provision of supports should take into account the demands of a given context (environment) and the individual's strengths, limitations, and preferences in that environment when determining what supports would allow the individual to be more successful in his or her daily life environments.

This new classification system focuses on assessing areas of needed supports and the intensity, frequency, and duration with which the supports should be provided. In 1992, Luckasson et al. described four levels of support intensity within this system: (a) intermittent, which is available when needed but not necessarily required all of the time; (b) limited, which may be needed over a defined period of time but not generally across all settings or across the lifespan; (c) extensive, which is needed regularly across most environments in which the individual functions and will likely be needed across the life span; and (d) pervasive, which is intense, required across all settings in which the individual is involved, and will be needed throughout the person's life.

Causes

The etiology of mental retardation is complex because it is often the result of interrelated risk factors. For example, a child found to have mental retardation may have been born to a teenage mother living in poverty who engaged in drug and alcohol use during the early stages of her pregnancy. This scenario includes several risk factors: a very young mother who may not be physically or emotionally ready for motherhood, possible inadequate maternal nutrition and prenatal care due to poverty, and substance abuse. It is probable that the combination of these risk factors affected the outcome for her baby, and it would be difficult to single out one cause for her child's disability.

The AAMR/AAIDD uses a multifactorial approach to determining risk factors that may result in intellectual disability. The categories of risk factors are as follows:

- Biomedical, such as metabolic disorders;
- Social, such as lack of neonatal care;
- Behavioral, such as parent substance abuse; and
- Educational, such as inadequate or delayed early intervention services.

Risk factors can be further analyzed in relation to when they occur (e.g., before birth, around the time of birth, or well after birth). A multifactorial conceptualization of the etiology of intellectual disability recognizes the complexity of conditions resulting in mental retardation and is useful in working toward its prevention. This approach also recognizes that addressing risk factors may not always focus solely on the child with the disability. Rather, effective prevention may focus on the child's family and society because some factors affect parents and families as well as children. To return to the previous example of interrelated risk factors, providing affordable prenatal care to a teenage mother living in poverty, assisting her to find safe housing and access to nutritious food for herself and her child, and providing postnatal follow-up care can be highly effective ways for a society to prevent conditions resulting in intellectual disability.

When considering various etiologies of mental retardation, a few generalizations can be made. In many instances, especially in the higher IQ ranges of mental retardation, a single, specific causal factor for the intellectual disability cannot be determined. For individuals at the lower IQ range of mental retardation, a causal factor can be identified approximately 60%–75% of the time and is most often related to a biological risk factor, such as a prenatal infection. Recent scientific advances have increased the number and precision of diagnoses of syndromes associated with intellectual disability (e.g., Angelman syndrome). New technology has also led to the discovery of previously unknown, genetically based syndromes, and it is likely that continuing medical diagnostic advancements will lead to identification of additional biological risk factors.

It is important to remember that even when there is a clearly identifiable biological causal factor, other risk factors may also influence the development, and, ultimately, the functioning of a particular individual. Knowing the etiology of the disability can provide much useful information; however, persons with the same etiology may be very unlike one another. Down syndrome is an example of this. Individuals with Down syndrome have a genetic difference that causes physical changes in their bodies, including their neural systems. Yet no two people with Down syndrome will be exactly alike, or function in exactly the same manner. These differences are the result of the interaction between their physical and neurological makeup and the social, educational, and behavioral factors present in their environments. Professionals would be wise to remember the old saying that "Diagnosis is not prognosis" and to consider each person with mental retardation individually.

Another issue to consider is that intellectual disability can occur with other mental and/or physical disabilities. Individuals with mental retardation may, for example, have a seizure disorder, experience sensory and motor impairments, or have a mental disorder. Consideration of these accompanying conditions is critical when planning educational, social, behavioral, and medical supports.

Future Directions

As our understanding of the nature of human functioning has deepened and our expertise in providing educational, social, and behavioral supports has increased, our expectations for persons with mental retardation have expanded. Attitudes toward and services provided for persons with mental retardation have undergone many positive changes, particularly in the past 25 years. Individuals with this disability are increasingly being successfully educated within inclusive settings with their age peers and are demonstrating the ability to acquire knowledge and skills beyond what many professionals considered possible. New vocational opportunities, more flexible service delivery systems, and more skilled provision of individualized supports have allowed many persons with mental retardation to obtain meaningful work, live in their own homes, be active members of their families and communities, and create personally satisfying lives for themselves.

Susan R. Copeland and Ruth Luckasson

See also Disabilities; Intelligence and Intellectual Development; Risk Factors and Development; Special Education

Further Readings

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., Text rev.). Washington, DC: Author.

Edgerton, R. (1967). *The cloak of competence.* Berkeley: University of California Press.

Luckasson, R., Borthwick-Duffy, S., Buntinx, W. H. E., Coulter, D. L., Craig, E. M., Reeve, A., et al. (2002). *Mental retardation: Definition, classification, and systems of supports* (10th ed.). Washington, DC: American Association on Mental Retardation.

Nirje, B. (1969). The normalization principle and its human management implications. In R. Kugel & W. Wolfensberger (Eds.), *Changing patterns in residential services for the mentally retarded* (pp. 181–194). Washington, DC: President's Committee on Mental Retardation.

Schalock, R. L., & Luckasson, R. (2005). *Clinical judgment.* Washington, DC: American Association on Mental Retardation.

Sparrow, S., Balla, D., & Cicchetti, D. (1984). *Vineland Adaptive Behavior Scales.* Circle Pines, MN: American Guidance Service.

World Health Organization. (2001). *International statistical classification of diseases and related health problems* (10th ed.). Geneva: Author

Web Sites

American Association on Intellectual and Developmental
Disabilities (formerly the American Association on
Mental Retardation): http://aamr.org

Meta-Analysis

A single study conducted to examine an issue will not usually establish definitive conclusions. It takes the accumulation of results across studies to begin to establish facts that can be then used to either validate a theory or formulate a new one. For example, many studies have been conducted to examine the relationship between the theoretical concept self-efficacy and academic achievement. That is, does a person's belief in his or her ability to perform academic tasks such as taking tests and completing homework assignments have an impact on academic achievement (e.g., grades, test scores)?

The traditional way to accumulate knowledge across these studies was to conduct a review of the literature. A researcher read all the published studies on the topic and then wrote a narrative describing these studies and an overall set of conclusions. Although these reviews were useful in organizing the studies conducted on a certain topic, the conclusions were based on the subjective impressions of the reviewer. Other concerns were the uneven quality of the studies as well as the difficulty in summarizing some topics in which there were very large numbers of empirical studies, often with what seemed to be conflicting results.

Meta-analysis is an attempt to address the weaknesses of the traditional literature review of empirical studies by using statistical integration and analysis of research findings. Data from studies examining the issue of interest are collected and aggregated, and then statistical tests are conducted to the aggregated or pooled data for the researcher to interpret. Thus, the primary purpose of meta-analysis is two fold, to first summarize the results of empirical research studies and, second, to estimate what the results might have been if all the relevant studies had been conducted without methodological limitations. This second purpose is expected to better reveal the underlying construct-level relationships in which scientists are most interested.

Data are quantified in two important ways in a meta-analysis: (1) The descriptive data from each study are coded; and (2) the results of each study are transformed into an effect size, which is a common metric across studies. This common metric transformation permits the data from different studies to be aggregated and compared, with effect sizes generally weighted to give more emphasis to studies using more participants. Effect sizes are largely sorted into two main types, correlation and standardized mean difference. The product-moment correlation coefficient and its variants provide an overall estimate of the strength of the relationship between two variables. If the focus of the meta-analysis is the effectiveness of some type of treatment or program, then standardized group mean differences are computed to provide an index of effect (e.g., Cohen's d). Research questions for the latter type of explanatory meta-analysis are often in relative terms, such as determining if one type of intervention is better than another type for this particular problem.

In the meta-analysis examining the relationship between self-efficacy and academic achievement, Karen Multon and her colleagues found an overall moderate effect size across all studies, which meant that a significant positive relationship was found between one's beliefs in one's ability to perform academic tasks and one's academic achievement. Additional statistical analysis showed that the variance in reported effect sizes was partially explained by certain study characteristics. For example, there was a stronger overall effect size for low-achieving students than for those students making normative academic progress. Thus, a meta-analysis is conducted not only to compute an overall effect size, but also to examine the relationship between the dependent variable (i.e., effect size) drawn from each study and the independent variables or characteristics of each study (e.g., population attributes, outcome measures, intervention).

An important benefit of meta-analysis is that it can be used to correct for measurement error, sampling error, and other artifacts that may distort study results. Essentially, meta-analysis can be used to estimate with greater accuracy the population parameters or values that would be obtained if one were able to conduct a perfectly designed study or sequence of studies with the entire population of interest. By synthesizing the results of independent research reports, a systematic error in one study will have less of an impact on the overall results. Another advantage of meta-analysis is that it combines knowledge succinctly no matter how many studies have been conducted on the topic, in contrast to a narrative review, in which studies may be

selective. This concise format aids in the distribution of information on the topic, as demonstrated by text-book authors heavily incorporating meta-analytic finds to make their manuals more accurate. Furthermore, meta-analysis may highlight particular gaps in the literature that can provide direction for future research.

Because the success of meta-analysis relies upon other studies, the failure of meta-analysis also lies in these same studies. The strength of meta-analysis is dependent on the quality of the studies included, incorporating studies that may be incomplete, have significant error, or have dissimilar research objectives. Since only studies that are published are used in a meta-analysis, the technique is inherently biased in that unpublished studies that did not achieve significant results are not included, thus leading to a potential inaccurate effect size estimate. Although the power of meta-analysis lies in the generalization and strengthening of a relationship, the method sacrifices the detailed information (e.g., finding a positive relationship with test scores, but the quantity is unknown). A more detailed meta-analysis can also be conducted for various subgroups (e.g., gender), but the researcher must undertake caution when interpreting the results because inappropriate subgroups can be inadvertently created (e.g., mixture of control and treated groups).

Karen D. Multon and Jill S. M. Coleman

See also Correlation; Inferential Statistics; Self-Efficacy

Further Readings

Glass, G. V, McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research*. Beverly Hills, CA: Sage.

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings* (2nd ed.). Thousand Oaks, CA: Sage.

Multon, K. D., Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology*, *38*, 30–38.

METACOGNITION AND LEARNING

Psychologists interested in how individuals learn have devoted much attention to the cognitive processes involved in encoding, storing, and retrieving information of all types, as well as comprehending the complex information with which they are confronted daily. Investigators have examined a variety of cognitive processes, with particular attention to memory and language comprehension. Such investigations have led to an understanding of the factors that lead to enhanced comprehension and recall during learning.

During the past couple of decades, researchers have also gotten keenly interested in metacognition. The term *metacognition* refers to the knowledge of *and* monitoring of cognitive processes. Because there is more than one cognitive process involved in learning, it is not surprising that researchers use more specific terms to denote the knowledge of and monitoring of different cognitive processes. For example the terms *metamemory* and *metacomprehension* are used to refer to individuals' knowledge of and monitoring of memory and comprehension, respectively. Most research on metacognition has been on metamemory or metacomprehension, although the metacognitive processes involved in performing other tasks, such as problem solving, have also been studied. Additionally, researchers have begun to explore metacognition outside of the laboratory, extending research paradigms to the classroom and other applied settings. There has also been an increase in attention paid to the role of social influences on metacognition.

Although the literatures on metamemory and metacomprehension are similar in many ways (e.g., the issues investigators are examining in the two literatures have much in common, and there are some similarities in the research paradigms employed), researchers examining metamemory have tended to use lists of isolated words as learning materials, whereas researchers examining metacomprehension have tended to use texts as learning materials. The present entry focuses on the role of metacomprehension in learning, as the text materials used in metacomprehension research are quite similar to the types of information typically encountered in learning in the classroom as well as other real-world settings.

Knowledge About Cognition

As individuals develop, they accumulate a great deal of knowledge as a result of life experiences. This knowledge can be thought of as "knowing that" knowledge (for example, knowing that a dog is a type of animal), also referred to as *declarative knowledge*, or "knowing how" knowledge (for example, knowing the procedures involved in typing), referred to as

procedural knowledge. One of the types of declarative knowledge that individuals acquire is knowledge about their own and others' cognitive processes.

Psychologists have primarily studied three components of metacognitive knowledge. These are *person* knowledge, *task* knowledge, and *strategy* knowledge. An example of person knowledge would be knowing that sixth graders are more likely to understand information about global warming than are second graders. An example of task knowledge would be the knowledge that it is easier to understand a passage when one is familiar with the topic than when one is unfamiliar. And, finally, an example of strategy knowledge would be the knowledge that rereading is a useful strategy when confronted with material not understood.

People acquire knowledge of cognitive processes, including person, task, and strategy knowledge, from a variety of sources. An individual may receive some instruction about person, task, and strategy knowledge from parents and teachers, but may primarily obtain such knowledge through a multitude of personal and informal learning experiences, including the observations of others as they attempt to learn.

Assessing Knowledge About Comprehension

The most direct way to assess individuals' knowledge about comprehension is through questionnaires or interviews. Several questionnaires have been developed to assess individuals' knowledge about comprehension, including person, task, and strategy knowledge, and questionnaires about comprehension have been developed to examine knowledge in children as well as adults. Although the data gathered from such questionnaires are interesting in their own right (for example, to reveal developmental changes in children's or adults' knowledge), what is of most interest to researchers is the relationship between individuals' knowledge and their performance on comprehension tasks or academic achievement.

Research on knowledge of comprehension in children reveals that such knowledge is related to both children's reading ability and age. Although these results are not terribly unexpected, an interesting finding in the literature is that adults' knowledge of comprehension processes has also been found to predict their comprehension performance and academic achievement. Thus, adults who have acquired more

general knowledge about comprehension (regardless of the source of that knowledge) have been found to perform better on measures of comprehension and to be more successful academically. Some researchers have found that strategy knowledge, in particular, is consistently related to comprehension performance. Thus, systematic rather than haphazard instruction in comprehension processes may be helpful to students.

Monitoring Cognition

During learning, it is important for individuals to both assess how well they are doing on a task, and initiate a plan to correct any problems they may be experiencing. These combined activities are referred to as *monitoring* cognition. Thus, while one is listening to a lecture or reading a book, it is important to both evaluate one's level of understanding (with perhaps a simple question such as "Do I understand what has been said/read up to this point?") as well as regulate understanding with one or more strategies if one is aware of a comprehension difficulty (for example, asking a question of an instructor or rereading a section of a paragraph not understood). Thus, monitoring of cognition really has two components. The first is *evaluation* of progress toward a cognitive goal, and the second is a *regulation* of activities through the use of appropriate strategies. If a student is regulating his or her cognition, then he or she has already attempted to evaluate progress. However, it is possible for a student to fail to evaluate progress, or also possible to evaluate progress (and find progress deficient in some way) but then fail to use one or more regulation strategies. Many fail to use strategies to repair comprehension problems simply because they lack the time or motivation to do so.

Unfortunately, the failure to evaluate one's progress and/or use strategies to aid progress toward goals is an all too common occurrence in students' efforts at learning. Furthermore, these types of monitoring activities (much like knowledge of comprehension processes) are often not taught directly, and, for this reason, students' learning and their ability to know how to learn may be hindered.

Assessing Monitoring of Comprehension

The majority of investigations of students' monitoring abilities have focused on students' ability to

evaluate rather than regulate comprehension during reading. Most investigators have conducted studies using college students as research participants, but some have examined children's ability to evaluate their comprehension.

Investigators have primarily used two research paradigms to examine students' ability to evaluate their comprehension during reading. One of these paradigms is the *error detection paradigm.* In this paradigm, students are given passages to read. Some of the passages contain an error, such as a nonsense word, false information, or a pair of inconsistent sentences. What is of interest is whether students notice the errors during reading. If not, investigators conclude that students may not be adequately evaluating their understanding during reading.

Linda Baker has argued that the ability to evaluate understanding during reading is not a unitary process, but rather is multidimensional, as individuals must really evaluate what they are reading using different standards of evaluation. Some standards of evaluation may be more difficult to use than others. Karen Zabrucky and DeWayne Moore, for example, found that children were better able to use a lexical standard or an external consistency standard (that is, they were better at evaluating their understanding of the individual words in a passage or whether the information fit with their own prior knowledge) than an internal consistency standard (evaluating whether information within a text was internally consistent). Researchers have generally found that children's ability to evaluate their comprehension, as measured by the error detection paradigm, develops with age. However, even college students frequently fail to use an internal consistency standard of evaluation during reading tasks. In fact, it appears as if evaluation skills continue to develop in college and graduate school, as students have more and more experience knowing how to learn.

Another, more widely used, research tool is the *calibration of comprehension paradigm.* In this paradigm, students are presented with several passages to read. Generally speaking, the passages are either unaltered or minimally altered, and they are obtained from textbooks or other reading materials. In the calibration of comprehension paradigm, students are asked to read each passage, one at a time, and provide ratings regarding their level of passage understanding or their readiness to be tested over the material. The similarity between the calibration of comprehension paradigm and the types of self-questioning activities in which students engage (or should engage) during everyday learning is strikingly apparent.

After students provide ratings of their understanding and/or test readiness, they are given a test over the passage information. What is of interest to psychologists is the relationship between students' ratings and their actual comprehension performance. This relationship is referred to as *calibration of comprehension.* Students who rate their comprehension high or indicate that they are ready for the test, and who perform well on the test, are said to be well calibrated. Similarly, students who rate their comprehension low or indicate that they are not ready for the test, and who perform poorly on the test, are also well calibrated (despite their poor comprehension performance!). Poor calibration is said to occur when there is a mismatch between one's self-assessment (or evaluation of understanding) and one's comprehension performance. The most common type of mismatch is an *illusion of knowing*, that is, believing that one understands something or is ready for a test when one is not. The implications of illusion of knowing for studying and learning are profound. If students exhibit an illusion of knowing, and research suggests that they frequently do, they will fail to continue the critical studying needed to understand and remember lecture or book material.

Several factors seem to influence calibration of comprehension. Students seem to be better able to calibrate their comprehension of text material when they are required to process a text more deeply, or when they reread passage information. Researchers are continuing to examine factors that are related to and may improve calibration ability. Research findings suggest that providing students with tasks during reading that contribute to more thoughtful and deeper processing would be highly beneficial. Also, practice at calibrating (for example, requiring students to assess their readiness and then providing them feedback on actual performance) might be helpful in reducing the illusion of knowing that so often accompanies students' decisions to discontinue their efforts during learning.

Karen M. Zabrucky and Lin-Miao L. Agler

See also Cognitive View of Learning; Reading Comprehension Strategies

Further Readings

Baker, L. (1985). How do we know when we don't understand? Standards for evaluating comprehension.

In D. L. Forest, G. E. MacKinnon, & T. G. Waller (Eds.), *Metacognition, cognition, and human performance* (pp. 155–205). New York: Academic Press.

Karabenick, S. A. (1996). Social influences on metacognition: Effects of colearner questioning on comprehension monitoring. *Journal of Educational Psychology*, *88*(4), 689–703.

Lin, L., & Zabrucky, K. M. (1998). Calibration of comprehension: Research and implications for education and instruction. *Contemporary Educational Psychology*, *23*, 345–391.

Moore, D., Zabrucky, K., & Commander, N. E. (1997). Validation of the Metacomprehension Scale. *Contemporary Educational Psychology*, *22*, 457–471.

Nietfeld, J. L., Cao, L., & Osborne, J. W. (2005). Metacognitive monitoring accuracy and student performance in the postsecondary classroom. *Journal of Experimental Education*, *74*(1), 7–28.

Veenman, M. V. J., Van Hout-Wolters, B. H. A. M., & Afflerbach, P. (2006). Metacognition and learning: Conceptual and methodological considerations. *Metacognition and Learning*, *1*, 3–14.

Zabrucky, K., & Moore, D. (1989). Children's ability to use three standards to evaluate their comprehension of text. *Reading Research Quarterly*, *24*, 336–352.

Mnemonics

Remembering information when one needs it, whether it is the name of an old friend one meets on the street, the answer to a question in a game of Trivial Pursuit, or the facts needed to answer a question on an exam is a challenge all people have faced. *Mnemonics* are aids or techniques for organizing information and encoding information to help an individual recall it when he or she needs it. The word *mnemonic* is derived from the ancient Greek word *mnemonikos* ("of memory"), and the first reported use of mnemonics appears in Cicero's *De Oratore.* In Cicero's account, the poet Simonides was dining with many other guests in the home of a wealthy merchant after having recited one of his poems. During the affair, Simonides was summoned outside the banquet hall. While Simonides was outside, the roof collapsed, crushing all the guests beyond recognition, but Simonides was able to identify all the bodies by remembering where each person had been sitting. Simonides' technique came to be known as the *method of loci*, and it has proven to be widely applicable.

In the method of loci, an individual associates the information he or she wants to remember with a series of familiar locations, such as rooms in the individual's house or landmarks on his or her route to work. For example, when arriving at home, you might park in the garage, enter through the laundry room, proceed through the kitchen, go through your bedroom to the closet, and so forth. You use this sequence by forming a mental image of each location in the sequence and of each item of information to be remembered, and then "placing" the images for the information you want to remember in one of the locations. Later, you can retrieve the information you wanted by retracing your path, stopping at each location to pick up the needed information.

Although an individual may not be familiar with the term *mnemonics* or the method of loci, it is almost certain that he or she uses a number of mnemonic devices. For example, if you want to remember the number of days in July, you are likely to use the familiar poem, "Thirty days hath September, . . ." Your spelling may have been improved by learning "*i* before *e* except after *c*." Many people learned their letters with the help of the alphabet song. A teacher may have helped you to distinguish between stalactites and stalagmites by pointing out that stalactites hang *tight* to the ceiling and stalagmites need all their *might* to grow up from the ground. If you ever wanted to be able to remember the first 15 digits of pi, you might try learning the following sentence: "Yes, I need a drink, alcoholic, of course, after the heavy sessions involving quantum mechanics." Once you are able to remember the sentence, you simply count the number of letters in order (e.g., Yes I need = 3.14). As you can see, the use of mnemonics is pervasive, and mnemonic devices take many forms.

Acronyms are words (or approximations to words) formed by the first letters of the items to be remembered. For example, if you can remember the order of colors in a rainbow, you may also be familiar with ROY G. BIV. If you have trouble remembering the names of the great lakes, you might try using the word HOMES (i.e., Huron, Ontario, Michigan, Erie, Superior). Acronyms frequently are used to make the names of organizations and other information more memorable (e.g., NATO, for Northern Atlantic Treaty Organization; AIDS, for acquired immunodeficiency syndrome). You may also have used *acrostics*, a closely related type of mnemonic in which the first letters of words in a sentence or phrase are used to cue memory. For example, if you know how to read music, you may have learned the names of the lines in the treble clef

with the assistance of "*E*very *g*ood *b*oy *d*oes *f*ine." Back when the solar system had nine planets, a number of sayings were used to help remember them in order of their distance from the sun, including, "*M*y *v*ery *e*legant *m*other *j*ust *s*at *u*pon *n*ine *p*orcupines."

Several types of mnemonics have been extensively studied by educational and cognitive psychologists. One such system, which has been proven effective in experiments in which people must recall a list of words in order, is the *peg word* method. In the peg word method, a highly imageable word that rhymes with a number is used as a retrieval cue. A commonly used peg word list begins as follows:

One is a bun.

Two is a shoe.

Three is a tree.

If an individual needed to remember a list of words in order (a common task in memory experiments of the mid-1900s), first he or she would memorize the list of peg words. Then, the individual would form a vivid mental image in which the first word in the list he or she is studying and the first peg word interact in some memorable fashion (e.g., if the first word were *deck*, an image of a hamburger bun sitting next to a plate made of playing cards, or a poker player with a sandwich in one hand and playing cards in the other). This process is then repeated for subsequent words from the list and peg words. To recall the words, the individual would count off the peg words, generate the associated images, and retrieve the words he or she was supposed to remember.

Unlike the loci and peg word systems, keyword mnemonics are custom-made for the information to be recalled. A *keyword* is a familiar, highly imageable word that sounds like all or part of the word to be remembered. Keywords were originally used to teach foreign vocabulary. For example, to learn the Spanish word for *letter* (of correspondence), *carta*, you could use the keyword *cart* and form an interactive image of an oversized letter in a shopping cart. Then, when asked to recall the Spanish equivalent of *letter*, you would recall the image and tranform the keyword to the proper Spanish term. The keyword method also has been used to help students remember other material, such as state capitals. Suppose you needed to know the capital of Maryland (Annapolis). Using the keywords *apple* and *marry*, you could form an image of two apples getting married.

People can be trained to generate keywords and images for themselves, but young children and people with cognitive impairments have trouble doing so. They can still benefit, however, when provided with the keywords and pictures. Although most experimental studies of keywords have demonstrated their effectiveness, some suggest that the use of keywords may not lead to improved long-term retention under all conditions. For the peg word and keyword mnemonics to be effective, the images must be vivid and interactive. The use of bizarre images has been highly touted, but research has failed to prove that they are any more effective. Because of their reliance on imagery, the use of these mnemonics to remember abstract words (e.g., justice, competence) tends to be more difficult and less effective.

As the examples above indicate, most mnemonics of all types are not meaningfully connected to the information they help people recall; therefore, their use has been criticized as artificial. In addition, most actually increase the amount of information that an individual has to remember (e.g., the sentence used to remember the value of pi). Nevertheless, through their use of imagery, rhyme, and familiar or meaningful material, mnemonics have been proven effective under a wide variety of conditions. As long as people have difficulty recalling information, mnemonics are likely to remain in use. However, as the technology for ready access to externally stored information continues to advance (e.g., computers, Internet, personal digital assistants), our dependence on them may wane, continuing the trend that has accompanied the advent of writing and the printing press.

Ernest T. Goetz

See also Learning Strategies; Long-Term Memory; Memory

Further Readings

Bellezza, F. B. (1981). Mnemonic devices: Classification, characteristics, and criteria. *Review of Educational Research*, *51*, 247–275.

Cook, N. M. (1989). The applicability of verbal mnemonics for different populations: A review. *Applied Cognitive Psychology*, *3*(1), 3–22.

Higbee, K. L. (2001). *Your memory: How it works and how to improve it* (2nd ed.). New York: Marlowe & Company.

Levin, J. L. (1981). Keywords in the classroom: The mnemonic 80's. *Educational Psychologist*, *16*, 65–82.

Thomas, M. H., & Wang, A. Y. (1996). Learning by the keyword mnemonic: Looking for long-term benefits. *Journal of Experimental Psychology: Applied*, *2*, 330–342.

MODE

The most general and least precise measure of central tendency is the mode. It is the value that occurs most frequently.

To compute the mode, follow these steps.

1. List all the values in a distribution, but list each only once.
2. Tally the number of times that each value occurs.
3. The value that occurs most often is the mode.

For example, if one were to examine different styles of learning and categorize learners into Types 1, 2, and 3, the results might be as shown in Table 1.

Table 1   Sample Data: Learning Style Frequency

Learning Style	Number or Frequency
Type 1	28
Type 2	53
Type 3	110

The mode is the value that occurs most frequently, which in the above example is a Type 3 learning style.

The most commonly made mistake when computing the mode is when the number of times a category occurs is selected rather than the label of the category itself. Instead of the mode being Type 3, it is easy to conclude the mode is 110. This is because one is looking at the number of times the value occurred, and not the value that occurred most often.

Distributions of scores can have more than one mode, such as the case where two categories of events occur a similar number of times, making the set bimodal in nature.

The mode should be used when the data are categorical in nature and values can fit into only one class, such as learning style, school attended, year in school, and political affiliation.

Neil J. Salkind

See also Mean; Median

Further Readings

Salkind. N. (2004). *Statistics for people who (think they) hate statistics* (2nd ed.). Thousand Oaks, CA: Sage.

MONTESSORI SCHOOLS

Montessori schools are learning communities that adhere to the educational approach developed by Maria Montessori. Montessori is both a pedagogical method and an international movement for peaceful, child-centered schooling, with more than 5,000 schools in the United States and another 2,000 in 70 other countries around the globe. Because the earliest Montessori schools served pre-school-aged children, Montessori is often associated with early childhood education. However, the Montessori approach spans the developmental continuum, with many schools serving children from infancy through adolescence and some through high school. Many of Montessori's claims relative to human development have been validated by experimental psychology. Similarly, her key educational innovations, such as the use of manipulative materials, child-sized furniture, and differentiated instruction, have been incorporated into mainstream classrooms.

History

Montessori education began in 1907 when Maria Montessori, one of Italy's first female physicians, opened the first Casa dei Bambini as part of an urban renewal project in the San Lorenzo district of Rome. In her early years as a physician, Montessori held an appointment at the university hospital while also operating a private practice. As she practiced medicine among the poor in Rome, she was drawn to the condition of those children and youth who were called "feebleminded" or "deficient." The work of the French physicians and psychologists Jean-Marc Gaspard Itard and Edouard Séguin was particularly influential. Both devoted their careers to working with people with disabilities, and it was from this orientation that Montessori launched her own work with children.

From Itard, Montessori adopted the practice of studying children's activity in their environment and then adjusting the environment based on those observations. Séguin, a student of Itard's, had begun

developing instructional apparatuses specifically for mentally impaired children. In addition to a focus on didactic materials, Montessori also adopted Séguin's developmental orientation toward learning environments. That is, Séguin observed that the environment itself should be customized to the needs of children at various stages of development. In Montessori's hands, the notion of developmentally responsive environment, filled with carefully constructed didactic materials, became the "prepared environment," one of the cornerstones of the Montessori Method.

Soon after the opening of the first Casa dei Bambini, Dr. Montessori began to receive acclaim for the method. In 1909, following the opening of subsequent Casas and confident that the approach would hold universal appeal, Montessori published *Il Metodo della Pedagogia Scientifica applicato all'educazione infantile nelle Case dei Bambini*, whose English translation is simplified to *The Montessori Method.*

Montessori made two heavily promoted trips to the United States. During both of these visits, she lectured to sold-out auditoriums and in the process developed an ardent American following, which led to a rapid proliferation of American Montessori schools and societies. American interest in Montessori, however, was not all positive. From the beginning, critics on both sides of the Atlantic attacked Montessori education. One of the most vocal detractors was famed Teachers College professor and self-proclaimed follower of John Dewey, William Heard Kilpatrick. In 1914, Kilpatrick produced a scathing critique that railed against the "fallacies" of self-correcting materials, "outworn and castoff" psychological theory, and a sharp focus on concentration at the supposed expense of social development. In a similar vein, Charlotte Mason, a leader of the British infant school movement, lamented what she viewed as Montessori's overemphasis on academic learning at the expense of the play spirit necessary for a happy early childhood.

The combined forces of World War I and the critiques, which became persistent, brought about the near disappearance of the Montessori method from the American educational scene. Meanwhile, Dr. Montessori continued to develop the method, eventually producing "advanced" educational materials aimed toward elementary-age children. In 1929, she founded the Association Montessori Internationale (AMI) in order to protect the integrity of her ideas and supervise the spread of the method through training programs, pedagogical guidelines, and the support of affiliated national societies. At her death, in 1952, control of AMI passed to her only child, Mario Montessori, who went on to lead the movement until his own death in 1982.

At about the time of Dr. Montessori's death, a young American woman discovered Montessori education and began a correspondence with Mario Montessori. Nancy McCormick Rambusch, who is widely credited with bringing about the American revival of Montessori, took Montessori training in London; established a Montessori learning environment in her home; and, by 1958, became instrumental in founding the Whitby School in Greenwich, Connecticut. Soon after the founding of Whitby, Rambusch was appointed by Montessori to be the representative of AMI in the United States. With Montessori's approval, she went on to found the American Montessori Society (AMS), which for 3 years functioned as the U.S. branch of the AMI.

Early on in the relationship, however, tensions began to emerge between AMI and AMS leadership. Where the European leaders of AMI worked diligently to preserve the integrity of the method as Dr. Montessori had designed it, claiming that child development occurred in relatively predictable and universal patterns, the Americans insisted that Montessori must be adapted to the unique cultural and social context of American society. By 1963, tensions between these two wings of the movement became untenable, and AMI and AMS broke ties, leaving a legacy of philosophical and practical schism that continues to shape the movement.

Montessori schools and training centers proliferated throughout the 1960s, and beginning in the 1970s, public Montessori schools were incorporated into several district desegregation plans. By the turn of the 21st century, public Montessori programs had become a significant growth sector of the movement. Now boasting a time-tested approach to educational reform, many districts began offering Montessori options through charters, conversions, and magnets.

The Method

The Montessori Method is a developmental approach to education grounded in ongoing clinical observation of the maturing child. In a sequence that foreshadows Jean Piaget's framework of staged development (sensori-motor, pre-operational, concrete operational and formal operational), Montessori outlined four

"planes" of development. Like Piaget's stages, the planes map a progression from reflexive motor activity to concrete and abstract thinking. The child at 3, for instance, is said to be in the period of "absorbent mind." The child between 6 and 12 (what Montessori called "childhood") is distinguished by intensifying interest in imagination.

Montessori's planes run in 6-year cycles (birth to 6, 6 to 12, 12 to 18, 18 to 24), with most cycles subdivided into two distinct 3-year cycles. Moreover, within each developmental plane, but especially in the first plane, Montessori claimed that children pass through "sensitive periods" for particular intellectual, social, and moral awakenings. There are sensitive periods for language, movement, music, order, and so on. The central role of the adult is to recognize these sensitive periods and direct the child to work designed to foster those awakenings.

What Montessori called the "prepared" adult, in fact, is one element of a three-pronged framework for the method. Along with an adult whose preparation emphasizes clinical child study, Montessori education also requires a "prepared environment," filled with carefully organized, scientifically developed learning materials. There are materials for the development of sensorial discrimination, language, mathematics, geography, and practical life. Advanced materials (those designed for the elementary child) include special exercises in grammar, biological classification, geometry, and advanced mathematical operations. The properly prepared teacher is an adult who has been trained not only in the purpose and proper use of these materials, but also in the correct manner of interacting with children so as to "direct" the child toward the optimal use of materials at optimal developmental moments. The third element in the framework is the child who is free to work within the prepared environment.

Although prepared environments are distinctive depending on the developmental level, several aspects of their design are consistent. Key design elements were developed through experimentation. For instance, one feature of all Montessori learning environments is the orderly placement of materials on low, open shelves. Children are free to select materials from these shelves, perform the appropriate exercises independently, and return the work to its proper place. This self-directed cycle of work is a hallmark of Montessori learning, but the earliest Montessori environments did not begin with this assumption. Rather, Dr. Montessori made the innovation in response to the children's

behavior. After arriving late one morning, one of Dr. Montessori's assistants noticed that the children had opened the cabinet where the materials were stored and, feeling free to take the materials themselves, began working independently. Observing that the children, left to their own devices, chose to work constructively, Dr. Montessori not only made the materials accessible to the children, but began to formulate a theory of education that radically challenged the foundations of current thought on childhood and child-rearing.

Instead of viewing children as animal-like creatures in need of the coercive control of adults, Montessori observed in children a natural desire to learn and a tendency to behave calmly and peacefully while in the midst of purposeful activity. Deep concentration, in fact, seemed to bring about a transformation in the child. Children who entered the Casa inattentive or rambunctious, after a period of sustained concentration, were able to sustain focus and became visibly more relaxed as a result of activity that was intrinsically satisfying. Development, she posited, was not just natural, it was a need in the young child that, when met, would render the child secure and happy.

The prepared environment, then, evolved into a learning space that provides both freedom and structure. Freedom comes in the form of liberty to move about, to select work based on interest, and to work individually or in a group. Structure comes from the care with which the environment is prepared to meet the developmental needs of the child. Experimental psychology now affirms many of Montessori's early observations: the importance of movement, the senses, order, and choice in learning, as well as the power of intrinsic versus extrinsic motivation. Achieving the correct balance of freedom and structure remains a delicate art form, practiced by experienced adults who are sensitive to the subtleties of individual development.

Types of Montessori Schools

Because the Montessori name is not protected by patents or trademarks, any school wishing to call itself a Montessori school may do so. In the course of the past century, two major types of Montessori schools have evolved: traditional and progressive, and distinguishing among types of Montessori schools most reliably revolves around the type of training received by teachers. Traditional Montessori teachers tend to adhere to a strict interpretation of the method as defined by

Dr. Montessori and her followers and are more likely to have been trained at AMI-affiliated training programs. Progressive Montessorians, by contrast, more often search for points of agreement between Montessori and mainstream American educational approaches. A wide array of associations align themselves with the progressive wing of the movement.

Jacqueline Cossentino

See also Intrinsic Versus Extrinsic Motivation; Piaget's Theory of Cognitive Development

Further Readings

Cossentino, J. (2005). Ritualizing expertise: A non-Montessorian look at the Montessori method. *American Journal of Education*, *111*(2), 211–244.

Cossentino, J. (2006). Big work: Goodness, vocation, and engagement in the Montessori method. *Curriculum Inquiry*, *36*(1), 63–93.

Kramer, R. (1988). *Maria Montessori: A biography.* Reading, MA: Addison Wesley.

Lillard, A. S. (2005). *Montessori: The science behind the genius.* New York: Oxford University Press.

Montessori, M. (1964). *The Montessori method.* New York: Schocken. (Original work published in 1909)

Rathunde, K., & Csikszentmihalyi, M. (2005). Middle school students' motivation and quality of experience: A comparison of Montessori and traditional school environments. *American Journal of Education*, *111*(3), 341–371.

MORAL DEVELOPMENT

Morality is an aspect of human development that is universal but—unlike language or motor development—differs markedly across individuals, cultures, and situations. Some people have a poorly developed moral sense, whereas others appear to have a hyperdeveloped moral sense. There are also major differences between cultures in what is considered moral. Furthermore, people who act highly moral in one situation may react with little regard for morality in another. These common observations make moral development both problematic and intriguing.

Research Strategy

The first problem in investigating morality psychologically is to decide strategically where to begin; that is,

what constitutes the core phenomenon or phenomena, the point of entry into the research field? Some investigators, such as social learning theorist Albert Bandura, have selected conduct or overt moral behavior as the strategic entry point. Psychodynamicists, such as Sigmund Freud and others, have focused on moral affect or emotion, especially guilt and shame. Martin Hoffman, in particular, has emphasized the role of empathy in moral development. A third group of investigators, initially Jean Piaget and later Lawrence Kohlberg, have rejected both foci as strategic entry points and instead focused—as most philosophers have—on moral judgments and reasoning. A fourth, more recent approach focuses on what are described as moral intuitions. There are arguments for and against each approach as a research strategy.

Moral Conduct

The argument for focusing on overt conduct lies both in its practical importance (as the "payoff" of studying morality) and on classic studies by Hugh Hartshorne and Mark May in the 1920s purporting to show that honesty is very situation-specific rather than a characteristic of persons. They reported that cheating by children in one situation did not predict cheating in another situation. However, a later reanalysis of their data showed that there was a common honesty factor, indicating a moral core, in addition to the situation-specific components. Also, years later, Kohlberg argued that knowing whether a person acted apparently in keeping with moral norms did not indicate *why* he or she acted so. By "why?" Kohlberg was *not* referring to some ultimate cause, but rather to the reasoning behind and meaning of the act for the actor. Furthermore, many developmental psychologists view the reliance on modeling, contingency, and reinforcement, as posited by social learning theory, inadequate to account for the developmental changes that regularly take place in children's moral thinking, feelings, and conduct.

Parenting and Guilt

Freud and other psychodynamics have focused on moral affect or emotions, especially guilt and shame. Freud's original theory depicted the essential core of morality, termed the *superego*, as resulting from defensive identification with parents as a way of resolving Oedipal conflicts. In this view, guilt is the child's

rivalry with and anger toward parents that is redirected against the self, and it also provides a way of substituting an identification with a loss of affective connection. Originally, the theory was heavily biological in its assumed causes, but as it came to dominate the American intellectual scene, it was transformed into an environmental theory, emphasizing how parenting influences superego strength as reflected in guilt. For example, an early study by Lois Hoffman and Herbert Saltzstein demonstrated a relationship between harsh, coercive parenting or *power assertion* and both weak guilt and a susceptibility to external influences, and an association between empathy-inducing discipline, called *induction*, with high guilt and a more fully internal moral sense. These were interpreted as showing that parenting influences moral development.

However, as Kohlberg commented to Saltzstein, what he concluded from these results was that "when all else fails, hit the kid." Richard Bell, in a seminal article, had already argued for the possibility of children's "givens," such as temperament, influencing parenting. Today, it is commonplace to recognize the two-way causation between parenting and the child's behavior and development. Nonetheless, most investigators would advocate as much use of reasoning and explanation and as little coercion as is possible given the cognitive and emotional development of the child.

A general problem with psychodynamic explanations of conscience is that they are too deterministic, positing fixedness around the time of Oedipal conflict that ignores all that goes on between that period of childhood and adolescence and adulthood. Jerome Kagan's criticism of what he calls *infantile determinism* also applies to superego theory.

More recently, Grazyna Kochanska and colleagues have conducted a series of very interesting longitudinal studies on parenting and children's moral development. For example, in one study, a relationship was demonstrated between children's compliance with adults and psychological discipline, *induction*, but only for those children who already had a certain degree of anxiety. Nazan Aksan and Kochanska have also reported findings that argue that conscience has two basic components: self-control and empathy, and that these components respond differently to different environmental and temperamental inputs. This might help explain why some guilt-prone individuals nonetheless repeatedly commit transgressions.

The distinction between shame and guilt has been a contentious issue for those who center their attention on moral emotions. Initially, the distinction between these two morally relevant affects was that shame was external, in the sense that it was triggered by the disapproval of others, whereas guilt was internal. The claim was further made by some anthropologists that certain cultures, especially Japanese and other Asian societies, are *shame cultures*, in which moral disapproval is dependent on the disapproval of others, whereas Western societies are *guilt cultures*, in which moral disapproval is independent of the reaction of others.

However, such neat distinctions may be problematic. For one thing, cultures are not equivalent to societies. For example, in large societies like the United States and Brazil, there are enormous regional differences. Also, socioeconomic class and ethnic differences often overwhelm societal differences. Indeed, urban-rural differences are often more profound than societal differences. In general, cultural differences are more subtle and variegated than these simple designations, such as *guilt* and *shame cultures* would imply.

The shame-guilt distinction has been transformed in other ways that help to shed light on the complexities of moral life. The psychoanalyst Gerhard Piers proposed that shame deals with failure to achieve a standard and the inevitable emotional abandonment that results, whereas guilt deals with aggression and the internal punishment that results. He also showed that the two moral emotions may be in conflict in ways that lead to apparently irrational and immoral behavior. A more recent, interesting distinction has been made by June Tangney and colleagues, who have proposed that guilt focuses the individual's attention on the victim of his or her transgression and thus may lead to constructive response, whereas shame focuses the individual's attention on the self and is more likely to be destructive.

Moral Judgment and Reasoning

The third approach to the psychological study of morality entered the scene with the publication of Piaget's study of *The Moral Judgment of the Child*. This was a detour for Piaget from his interest in general cognitive development. However, its influence is still being felt. The book is divided into a shorter section that describes the practices and beliefs of Swiss children about the popular games of marbles, and descriptions of children's judgments and reasoning about vignettes in which a fictive child committed an act that might be viewed as a moral transgression.

It has been suggested that in writing the book, Piaget may have been trying to counter the views of the great French sociologist Émile Durkheim, who posited obedience as the psychological bedrock of morality. (This idea has interesting, but largely unexplored, implications for conscience and mental health.) The function of obedience (to the group) was to curb individual appetites and thus help people avoid the despair that comes from unlimited goals. Piaget transformed this into a developmental form, with obedience and other nonmoral justification, which he called *heteronomy*, characterizing the thinking and conduct of young children, which was then replaced by *autonomy*. By this latter term, Piaget did not mean independence of others but rather an attitude of mutual respect, which he saw as arising primarily from peer interaction rather than parental influences.

Perhaps the best known of his findings was the gradual developmental shift in moral judgment from focusing on the outcome of an act to focusing on the intent motivating it, which he termed the shift from *objective to subjective responsibility*. Thus, a young child—roughly 5 to 8 years of age—would judge a child who accidentally breaks five dishes while trying to help her mother, as bad and worse than a child who mischievously breaks one dish or does so while trying to steal a cookie; an older child would judge the latter child as worse. This finding has sometimes been dismissed as due to some methodological confound (e.g., that the intent comes early in the story and thus is simply forgotten). Some believe that, despite these faults of method, this constitutes a real developmental shift not reduced to an artifact. Marla Johnston, a doctoral student working with Saltzstein, has shown that this shift from a focus on outcome to a more consistent focus on intent figures into how children view the fairness of parental discipline. She found that whereas older children (7–11 years) judge a mother as unfair who disapproves a well-intentioned act that results in damage and a mother who approves a bad-intentioned act that results in a good outcome, younger children (4–6 years) do not make the distinction. This shows that older children use their own moral judgments as a template against which to judge maternal intervention, but younger children do not.

Piaget's book, published in 1932, stimulated a flurry of interest that died down in the wake of the onslaught of psychodynamic and behaviorist theories. But this general interest in children's moral thinking was reawakened with the publication of articles based on Kohlberg's dissertation, which he described as "simply" to replicate and extend Piaget's theory, but went much beyond Piaget's theory and findings.

He presented children ranging in age from about 10 to 16 with relatively complex moral dilemmas, that is, choices between two moral duties rather than between a moral duty and a nonmoral need, and followed up their moral thinking longitudinally. The best known dilemma is named after its protagonist, Heinz:

> Heinz's wife is dying of a special form of cancer and only one drug can save her, but Heinz cannot afford the cost of the drug being sold by the pharmacist. Should Heinz steal the drug to save the wife's life or not steal it and let his wife die?

Rather than focus on what choice was recommended by the child, Kohlberg focused on the reasons the child gave for justifying whatever choice he or she advocated. By examining the child's reasons for justifying his or her choice, Kohlberg was treating the child as a young philosopher, which enabled him to scrutinize the structure or form of the reasoning. What is meant by the structure or form is not easy to convey; perhaps an example will help.

> What is the relationship between a duty and a right? If I have a duty to act in a certain way with regard to someone (spouse, child, friend, colleague, teacher, student . . .), does that other person have a right to expect me to do so? And vice versa?

Note that this example does not specify what the duty and right are, but rather their interdependent relationship. Accordingly, younger children may believe in an individual's right, but this has no implication for another's duty, whereas older children do not believe in an individual's right. (Rick Shweder has argued that in India and other non-Western cultures, duties and rights may not imply one another.) This characteristic of the concepts and coding of reasons is not easy to convey and take training and practice.

Although the focus of attention is usually on Kohlberg's stages, an interesting aspect of the theory is the assertion that moral beliefs and reasoning are neither copies of the environment (home, school, etc.) nor simply the unfolding of predetermined moral reactions, but rather qualitatively different forms (structures) of reasoning that emerge from attempts to

resolve social/moral conflict. For example, another of Kohlberg's dilemmas features conflict between a boy and his father: The father promised his son that he could keep his earnings from after-school work, but then reneged and demanded the money for the father's own fishing trip. On one hand, the son owes his father a lot for all that the father has done for him, but on the other hand, his father made a promise, and promises should be kept. According to Kohlberg, the qualitatively different reasoning that children of different developmental levels reveal show the increasing ability of these children to reconcile diverse perspectives and renders their reasoning increasingly complex and stable (or to use the Piagetian term, *equilibrated*). According to Piaget, Kohlberg, and other cognitive developmentalists, moral reasoning proceeds through a series of qualitative transformations in which the essence of morality justice or fairness becomes more and more separated from morally irrelevant concerns such as whether one gets punished or whether one's decision is approved by others. Their theories are based fundamentally on the moral philosophy of Immanuel Kant and neo-Kantians, such as John Rawls. This approach directs educators away from indoctrination and toward moral development as problem solving.

Gender Differences

In a much heralded study of women facing the question of abortion, Carol Gilligan and Mary Belenky argued that actual moral decisions, such as whether to abort a fetus, were much more nuanced, contextual, and attentive to consequences for significant others (often termed *caring*) than Kohlberg's focus on justice would allow. Sometimes, the claim has been made that women are guided by a concern for caring whereas men are guided by a concern for justice. This claim is almost certainly an oversimplification, and evidence shows that both men and women focus on justice and caring. Indeed, these two principles have a long history in philosophy, the former being an example of the *deontic principle*, with its focus on duty regardless of outcome, and the *beneficence principle*, with its focus on utilitarian concerns or consequences. Examples of the two are Kant's theory deontic and John Stuart Mill's utilitarian theory. These two general principles may not be reducible to a simpler one, and therefore conflicts between them may not be reconcilable.

More Recent Approaches

Another developmental psychologist, Elliot Turiel of Berkeley (who had originally worked with Kohlberg), has argued and demonstrated that young children (roughly by age 4 or 5) already distinguish different classes or *domains* of social rules: the *moral*, involving fairness or harm to others and which is understood to be universal and unchangeable, and *social conventions* or social regularities, which are understood to be particularistic rather than universal (e.g., a man removes his hat when entering a church, keeps it on when going into an orthodox synagogue, and removes his shoes when entering a mosque) and may be alterable by consent. A third category is the *personal* (such as which clothes an adolescent should wear), which is not theoretically subject to the claims of others and has been shown by Judy Smetana and others to be a particular arena of parent-adolescent conflict. This conceptualization rejects the idea that young children are *heteronomous*, that is, that they initially confuse what is genuinely moral from what is not (such as the physical damage done, punishment, etc.).

The dispute between Kohlberg's cognitive developmental approach and Turiel's domain approach has occupied many researchers and much journal space. Orlando Lourenço, a Portuguese psychologist, has argued that whereas Turiel and his colleagues have focused on simpler moral conflicts, such as whether it is right to breach some moral or social rule, Kohlberg focused on moral dilemmas, where two moral rules (e.g., not to steal but not to let your wife die) are in conflict. Thus, Turiel's findings may hold for simpler moral conflict, where a moral duty and a nonmoral urge are in conflict, and Kohlberg's for the more complex moral dilemma, where two compelling moral duties are in conflict.

Moral Thought and Action

One of the arguments against all cognitive approaches is the claim that the relationship between moral reasoning and conduct is weak and inconsistent. This claim goes back to the early findings of Hartshorne and May, who found zero-order relationships between moral beliefs and moral conduct in the form of cheating/honesty. However, unlike the ingenious methods for studying honesty, their measures of moral beliefs were simple and did not assess genuine moral reasoning. Moreover, a subsequent review of studies

by Augusto Blasi concluded the relationship between moral conduct and moral judgment/reasoning is spotty, weak, and inconsistent, depending on the particular aspect of moral behavior studied, but not nonexistent. The easy conclusion to draw is that moral talk is "cheap" and that moral reasoning is thus largely a justification of what has already been decided. In a theoretical analysis, Saltzstein argued that this conclusion is not warranted. His analysis rested basically on two points: First, measures of moral reasoning usually involved responses to complex moral dilemmas involving a clash between two duties (e.g., to be honest and not to steal and not to let your wife die), whereas the measures of moral conduct often involved simpler conflicts between a moral duty (e.g., not to cheat) and a nonmoral urge (e.g., to gain something illicitly). The second difference between moral judgments and moral conduct lay in the perspective taken by the respondent: that of an observer making a moral judgment of others versus an actor deciding on an action to be taken. This perspectival difference leads to different inferences about the "facts" of the situation. For example, it is well-known that observers often overlook the constraints of the situation and attribute the actor's behavior to the person, whereas actors see the constraining aspects of the situation and therefore attribute their behavior to these situational constraints. This and other differences in perspective may account for some of the differences found between moral conduct and judgment/reasoning. Thus, only a naïve person would claim that people regularly act on their moral beliefs, but it is also too simplistic to say that they *never* do.

Barbara Rogoff, a culturally oriented developmental psychologist at the University of California, Santa Cruz, has criticized approaches that emphasize reasoning, saying that they appear to assume that most of what we actually know can be articulated. She points out that an individual need not know formal grammar in order to actually use it to produce or understand language. Why make such an assumption about moral development? That is, moral decisions may not always be explainable, especially by children. In short, we may know more than we can say.

A radically different approach that does not rely solely on verbal explanation has been advocated by Jon Haidt. His model of moral behavior is complex and includes some of the cognitive factors on which developmentalists focus, such as reasoning. However, the central and novel part of his theory is that moral decisions are quick, intuitive, and involve little reasoning except as an afterthought. This approach is interesting and appears to solve the apparent problem that moral thought and conduct frequently go their own ways. Saltzstein and Tziporah Kasachkoff, along with others, have criticized Haidt's approach in print and in talks.

Morality and Culture

Cultures obviously differ in their moral beliefs and practices, and the study of variations in morality across cultures has been a popular and contentious area of study in terms of how to explain these differences and what their implications are for the universality or relativity of morality. People often proclaim that "you can't judge other cultures; they have their own unique sense of morality." The same individuals may express outrage at the genocide that is currently going on in Darfur and may belong to human rights organizations, which assume a universal set of moral rights and duties. How to reconcile this inconsistency?

The issue is complex, and persuasive arguments have been advanced by scholars taking different positions. Although Kohlberg usually argued that cultures or societies cannot and should not be rated on a moral development scale, individuals can, or more precisely, their reasoning can. Turiel has argued that the basic categories of domains (the moral, social-conventional, personal, and practical) are universal, although the practices and rules that are placed in these domains are specific to cultures.

Others, notably Shweder of the University of Chicago, have strongly advocated a different view, largely based on Shweder and his colleagues' (Joan Miller, Lene Jensen et al.) research in India. Shweder has advanced a view that stresses the uniqueness of each culture (e.g., Brahmin Indian in the state of Orissa) and the different emphases each culture gives to three different kinds of moral systems: *morality of rights, community*, and *divinity*. As an example, Brahmin widows are strongly enjoined from eating spicy food. What may appear to be a group-specific convention is not viewed as such by those who hold them. It has all the force of a moral stricture even though it does not seem to involve harm to others or violations of the principle of justice.

Turiel has argued that when one examines beliefs in specific cultural context, it becomes clear that the apparent cultural differences are not in morality as

much as in other aspects of thinking. For example, the dietary stricture mentioned above has a basis in the holders' beliefs about the world. He argues that when one understands their religious beliefs about reality (e.g., that if a widow eats spicy food or fish, it excites her and tends to make her more prone to sex, and this means that her husband will not find peace in the afterlife). If you held these religious beliefs, you would understand the harm done to others (deceased fathers-husbands) as they do! This is an important point. Nonetheless, the evidence of Shweder and colleagues does suggest that the social conventional-moral divide may not be universal.

Other cultural differences have been documented in moral and nonmoral aspects of social reasoning. For example, Joan Miller and her colleagues have shown that there are systematic cultural differences in (a) how morally relevant behaviors such as stealing or helping are explained differently in different cultures (e.g., generally attributed to external pressures in India but attributed to the individual in the West); (b) the degree to which optional and obligatory duties, and duty and caring, are distinguished; and (c) the degree to which exceptions are made depending on the specific context in which the moral decision is made (e.g., was it provoked, was the temptation great?). Furthermore, Miller has argued persuasively that culture needs to be considered intrinsic to explaining moral thought and practice, and not just considered an ''add-on'' or a modification of a basic, culture-free phenomenon.

One deficit of both theories (Turiel's and Shweder's) is their apparent omission of a cogent description of the developmental mechanisms for acquisition of moral beliefs, reasoning, categories, and so on. To state that these are learned through interaction (especially by observing consequences) and communication (by noting what adults emphasize) may be true but is insufficient to explain variations, especially systematic developmental changes, in moral thinking and practice.

Is there any common framework within which to compare such diverse cultures? Lene Jensen has offered a framework consisting of a set of universal questions that each culture must address (e.g., Why are we suffering?). These questions are what all cultures have in common. Their answers are where they differ.

Thus, the study of moral development leads to a number of critical issues, among them how human development relates to culture. Much has been clarified and learned by theorists and researchers, but much remains unclear, and what is unresolved clearly has important implications for education, public policy, and ethics.

Herbert D. Saltzstein

See also Cognitive Development and School Readiness; Kohlberg's Stages of Moral Development; Piaget's Theory of Cognitive Development

Further Readings

Arksan, N., & Kochanska, G. (2005). Conscience in childhood: Old questions, new answers. *Developmental Psychology, 41*, 506–516.

Bell, R. (1968). A reinterpretation of the direction of effects in studies of socialization. *Psychological Review, 75*, 81–95.

Freud, S. (1961). *The ego and the id* (J. Strachey, Trans. & Ed.). *The Standard Edition*, Vol. 19. (Original work published in 1923)

Gilligan, C. (1982). *In a different voice.* Cambridge, MA: Harvard University Press.

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgments. *Psychological Review, 108*, 814–834.

Kagan, J., & Lamb, S. (Eds.). (1987). *The emergence of morality in young children.* Chicago: University of Chicago Press.

Killen, M., & Hart, D. (1995). *Morality in everyday life: Developmental perspectives.* New York: Cambridge University Press.

Kohlberg, L. (1981). *Essays in moral development.* New York: Harper & Row.

Miller, J. G. (1996). Theoretical issues in cultural psychology. In J. W. Berry, Y. H. Poortinga, & J. Pandey (Eds.), *Handbook of cross-cultural psychology: Vol. 1. Theory and methods* (2nd ed., pp. 85–128). Boston: Allyn & Bacon.

Piers, G., & Singer, M. (1953). *Shame and guilt: A psychoanalytic study.* New York: International University Press.

Saltzstein, H. D., & Kasachkoff, T. (2004). Haidt's moral intuitionist theory: A psychological and philosophical critique. *Review of General Psychology, 8*, 273–281.

Turiel, E. (2002). *The culture of morality: Social development, context, and conflict.* New York: Cambridge University Press.

MOTIVATION

The English word *motivation* has its etymological roots in the Latin word *movere*, which means "to move." Although no consensual definition of motivation exists

in scientific psychology, the definitions provided by most theorists rest firmly on the notion of movement. The definition offered herein is no exception: *Motivation* is defined as the energization and direction of behavior. Motivation is important in educational psychology because it explains and predicts the behavior of students, teachers, and administrators in school settings. In this entry, the concept of motivation in general is overviewed, followed by an overview of motivation within the field of educational psychology and a final section on integrative models of motivation within educational psychology.

Motivation Conceptualized

Motivation is a hypothetical construct, meaning that it is an abstraction, not an overt entity that can be seen with the eyes. Although it is not observable in and of itself, motivation does have direct links to observable behavior. That is, motivation cannot be observed, but its effects may be observed. Specifically, one infers motivation from observing the movement of individuals. Such movement often appears to be systematic, rather than random, and the force that impels and guides this systematic movement is presumed to be motivation. For example, if one sees a father walking toward his daughter with a large smile on his face and his arms open wide, one may infer that the father is motivated to hug his daughter. Thus, motivation is conceptualized as a causal agent; it explains behavior; it does not simply describe behavior.

Motivation focuses primarily on two central questions—the *why* and the *how* of behavior. These questions map onto the energization and direction aspects of motivation, respectively. "Why" represents the underlying reason(s) that an individual is energized or impelled to engage in a certain type of behavior. This "why" question focuses on what the pioneering psychologist William James called the "springs of action"—the fundamental impetus for behavior that gets the individual oriented toward a certain type of movement. This description of energization does not assume that the individual is passive until instigated to action; on the contrary, people are viewed as perpetually active, with instigation functionally representing a shift from one form of orienting to another.

"How" represents the guiding or channeling of energization in a precise way. This "how" question focuses on the specific aims on which persons focus to direct their behavior. Both energization and direction,

why and how, need to be considered to fully explain motivated behavior. Accordingly, one may argue that motivational accounts of behavior must be hierarchical, in that they must articulate the different levels of motivation that operate in tandem to produce behavior.

The most basic distinction that can be made about motivation is whether it represents approach motivation or avoidance motivation. As 19th-century philosopher Arthur Shopenhauer observed, people are not simply motivated, they are motivated toward something or away from something. This approach-avoidance distinction is applicable to all types of motivation and to all types of organisms (e.g., from humans to the single-cell amoeba). Furthermore, the approach-avoidance distinction is applicable to both the energization and direction aspects of motivation. Approach motivation is the energization of behavior by, or the direction of behavior toward, positive stimuli (objects, events, possibilities), whereas avoidance motivation is the energization of behavior by, or the direction of behavior away from, negative stimuli (objects, events, possibilities).

The concept of movement is embedded within the approach-avoidance distinction, be it physical movement or psychological movement. Stimuli that are positively evaluated are inherently associated with an approach orientation to bring or keep the stimuli close to the individual (literally or figuratively), whereas stimuli that are negatively evaluated are inherently associated with an avoidance orientation to push or keep the stimuli away from the individual (literally or figuratively). Positively and negatively evaluated stimuli produce (at minimum) a physiological and somatic preparedness for physical movement toward and away from the stimuli, respectively, but this preparedness may or may not be translated directly into overt behavior. Furthermore, movement toward and away from a stimulus each can have two different forms. "Movement toward" can represent getting something positive that is currently absent, or it can represent keeping something positive that is currently present (functionally, continuing toward); "movement away" can represent keeping away from something negative that is currently absent (functionally, continuing away from), or it can represent getting away from something negative that is currently present. As such, approach motivation not only involves promoting new positive situations (e.g., striving to get a college degree), but it also involves maintaining and sustaining existing positive situations (e.g., striving to maintain an already high grade point average);

avoidance motivation not only involves preventing new negative situations (e.g., trying to avoid failing organic chemistry), but it also involves escaping from and rectifying existing negative situations (e.g., trying to move to a new apartment to get away from one's slovenly roommate).

Although motivation represents an internal force within the person, it is influenced by both internal factors within the person and external factors outside of the person. Broadly stated, internal factors include biologically based predispositions that lay the foundation for behavior across situations; affectively based and cognitively based dispositions that produce behavioral tendencies in particular domains; and situation-specific states that have an immediate, direct impact on behavior. External factors include culture, which provides a basic set of assumptions, meanings, and practices that establishes a person's basic worldview; socialization by parents, other adults in leadership positions, and peers that molds and shapes a person's specific values, beliefs, and behavioral patterns; and environmental contexts that provide immediate cues for what is important and expected in a given situation. Internal factors influence each other in producing behavior, and external factors likewise influence each other. Furthermore, internal and external factors influence external factors in influencing behavior, and vice versa. For example, external factors play an integral role in the development of many of the internal factors, but once in place, internal factors serve as a filter through which external factors are interpreted. In short, human motivation is complex, emerging from a multitude of mutually interacting internal and external factors.

Further highlighting the complexity of motivation is the fact that motivation can be conscious or unconscious. Although for many years, psychologists assumed that most, if not all, behavior is energized and directed by conscious processes, this assumption is no longer sustainable. Researchers have clearly documented that behavior is commonly a function of motivational tendencies that are activated automatically and directly produce behavior without the actor's conscious intention or even awareness. Indeed, it appears that conscious and unconscious aspects of motivation are often independent of each other, and that people's beliefs about the causes of their behavior are often dramatically off the mark.

Motivation in Educational Settings

In educational psychology, the type of motivation that is most central, and that has received the most theoretical and empirical attention, is competence motivation (often labeled "achievement motivation"). Competence may be defined as a condition or quality of effectiveness, ability, sufficiency, or success. Competence motivation may be defined as the energization and direction of competence-relevant behavior, or why and how people strive toward competence (success) and away from incompetence (failure). Applied to educational settings, competence motivation focuses most directly on how and why students strive toward success and away from failure in the classroom.

Research on competence motivation has a long and distinguished history in scientific psychology. Indeed, competence motivation has been the focus of theoretical and empirical work since the emergence of psychology as a scientific discipline in the mid- to late 1800s. The aforementioned James was one of the first in psychology to write about competence motivation, as he offered speculation regarding how competence strivings are linked to self-evaluation. A great deal of research has been conducted on competence motivation since the time of James, and it remains an extremely popular topic, both in psychology in general and in educational psychology in particular. Research on competence motivation is conducted both in the experimental laboratory, where variables are typically manipulated, and in real-world achievement situations such as the classroom, where variables are typically measured.

The aim of competence motivation research is to explain and predict any and all behavior that involves the concept of competence. It is important to note that the aim is *not* to explain and predict all behavior that takes place in achievement situations. A great deal of the behavior that takes place in achievement situations is not necessarily about competence per se. For example, in the classroom, students are not only motivated with regard to competence, but they are also often motivated with regard to relationships with their teachers and classmates, and with regard to instrumental concerns such as graduating to get a good job or to please their parents. These forms of motivation are important and clearly influence behavior in achievement situations. However, from the standpoint of competence

motivation, these types of motivation are primarily of interest to the extent that they have an influence on competence motivation. Importantly, in classroom contexts, these other forms of motivation do indeed have a strong influence on competence motivation, a fact that is not often acknowledged.

In order to study competence motivation, theorists and researchers create and use specific constructs that explain and predict behavior in competence-relevant situations. Many different competence motivation variables have been proposed and studied over the years, and in the following, the variables that have had the most central and enduring influence on the field will be briefly overviewed.

Intelligence and Ability

Intelligence and ability are competence itself and essentially represent the foundation on which competence motivation rests. Not surprisingly, competence is viewed as facilitating positive outcomes in achievement settings.

The traditional, and still quite popular, conception of intelligence/ability is that there is a general form of intelligence or general cognitive ability that is commonly labeled g. From this perspective, abilities in various domains are correlated with each other, reflecting a g factor that may be assessed with intelligence quotient (IQ) tests.

Other theorists advocate a different conception of intelligence/ability that focuses on the separability of a variety of more specific intelligences/abilities instead of (or, in some instances, in addition to) a unified g. For example, according to the triarchic theory of intelligence, intelligence/ability is best viewed in terms of three independent sets of abilities that help individuals adapt to real-world environments: analytical, creative, and practical. In similar fashion, the theory of multiple intelligences views intelligence/ability in terms of even distinct abilities that relate to different domains of activity: linguistic, musical, logical-metatheoretical, spatial, body-kinesthetic, intrapersonal, and interpersonal. Proponents of these differentiated models of intelligence/ability do not contend that all persons are of equal intelligence; rather, they emphasize that people have strengths and weaknesses in different areas. In addition, those adhering to a differentiated conception tend to view intelligence/ability as quite malleable and open to development, whereas those who hold a unified conception tend to view

intelligence/malleability as innate and relatively stable across the life span.

Perceived Competence

At least partially (and sometimes entirely) independent of actual competence is one's cognitive perception of one's competence. Perceived competence is relevant before, during, and after engagement in achievement tasks. Before a task, perceived competence represents a competence expectancy—one's expectation for how one will perform; during a task, perceived competence represents a competence perception—one's ongoing sense of how one is performing; and following a task, perceived competence represents perceived performance—how one feels one has performed. Each of these types of perceived competence has been shown to have an important influence on competence motivation, and in nearly all cases, perceived competence is viewed as facilitating positive outcomes in achievement settings.

Perceived competence may be examined at two basic levels, the ability level and the task level. With regard to the ability level, theorists talk of cognitively based self-perceptions of competence and posit different domains of competence perceptions, such as the academic domain and the social domain. Within each domain, these self-perceptions may be differentiated further, such that within the academic domain, for instance, one has a separable competence perception for math, science, English, and so on. Although theorists disagree on the precise way in which these various types of self-perceptions of competence are related to each other, it seems sensible to view them as hierarchically related, such that discipline-specific perceptions (e.g., perceptions of math competence) feed into domain-specific perceptions (e.g., perceptions of academic competence), which feed into an overall perception of competence.

Other theorists have taken a more situated, task-specific approach to competence perceptions. One important distinction that has been made in this regard is that between efficacy expectations and outcome expectations. An efficacy expectation is a personal judgment of one's ability to execute a particular behavior. An outcome expectation is one's personal judgment that a particular behavior, if successfully performed, will lead to a particular outcome. Research has shown that these two types of expectancies have a differential influence on achievement outcomes.

Achievement Motives

An achievement motive is a dispositional motivational tendency to energize competence-relevant behavior and orient individuals toward success or failure possibilities. Two primary achievement motives have been posited by theorists: need for achievement, which represents a desire to approach success, and fear of failure, which represents a desire to avoid failure. At the core of these motive dispositions is an affective sensitivity: Persons high in need for achievement experience a great deal of pride upon success, whereas persons high in fear of failure experience a great deal of shame upon failure. Individual differences in these affectively based motives are thought to be grounded in socialization, particularly early socialization experiences with one's parents involving rewards and punishments following demonstrations of competence or incompetence. In the main, need for achievement has been linked to optimal functioning in achievement settings, whereas fear of failure has been linked to maladaptive functioning.

Theorists distinguish between two general types of motives, including achievement motives. Implicit motives are rooted in midbrain structures and operate outside of conscious awareness, whereas explicit (or self-attributed) motives are rooted in cortical brain structures and are accessible to conscious awareness. Research has shown that implicit and explicit achievement motives, particularly need for achievement, are not related to each other very strongly. However, factors such as one's degree of access to bodily cues and one's level of preference for consistency greatly influence the extent to which implicit and explicit competence motives are related.

Achievement Values

An achievement value is the degree to which competence is considered important or valued by the individual. Theorists initially viewed achievement values in terms of affectively based incentives for success or failure. For example, the value of success was presumed to be high to the extent that a person anticipated experiencing a great deal of pride if success was achieved. Later theorists have retained the notion that value is grounded in affect, but have emphasized a general affective commitment to competence (labeled "competence valuation"), rather than anticipation of a specific affect upon success. Values are critical determinants of achievement behavior, as low

competence valuation may manifest as a divestment from competence altogether. Such divestment from competence strivings is sometimes referred to as *disidentification*, and it often occurs for defensive or self-protective reasons.

Some theorists have minimized emphasis on the affective basis of achievement values, opting instead to portray them as cognitive beliefs. From this perspective, values may be differentiated in terms of four components: Attainment value represents the personal importance attached to doing well on, or participating in, a given task. Intrinsic value represents the enjoyment one gains from doing the task itself or the anticipated enjoyment one expects to experience while doing the task. Utility value (i.e., usefulness) represents how a task fits into one's future plans. Cost represents how the decision to engage in one activity limits access to other activities. These four components of value are presumed to jointly contribute to the overall competence value a task has for an individual.

Achievement Goals

An achievement goal represents a concrete cognitive representation of a possible competence-relevant outcome that a person is committed to approach or avoid. These goals are viewed as situation-specific aims that establish a framework for how individuals engage in and experience achievement tasks.

Initial models of achievement goals posited that such goals may vary with regard to a performance-mastery dichotomy. Performance goals focus on demonstrating one's ability relative to others, whereas mastery goals focus on developing task mastery. More recent models have revised this dichotomous approach by incorporating the approach-avoidance distinction. One such model, the trichotomous achievement goal framework, identifies three distinct competence-based goals that one may adopt in achievement settings: mastery-approach goals (focused on attaining task-based or intrapersonal competence, sometimes simply labeled "mastery goals"); performance-approach goals (focused on attaining normative competence); and performance-avoidance goals (focused on avoiding normative incompetence). Another model, the 2×2 achievement goal framework, extended the trichotomous model by adding a fourth goal: mastery-avoidance goals (focused on avoiding task-based or intrapersonal incompetence). Generally speaking, the two approach goals have been shown to facilitate positive outcomes

(with mastery-approach goals being especially beneficial for phenomenologically based outcomes such as intrinsic motivation, and performance-approach goals being especially beneficial for performance-based outcomes). The two avoidance goals have been shown to lead to negative achievement outcomes, with performance-avoidance goals being particularly deleterious.

Implicit Theories of Ability

Implicit theories of ability refer to individuals' assumptions about the nature of competence and ability. The assumption that ability is a fixed and unchanging individual characteristic has been termed *entity theory*, whereas the assumption that ability is a dynamic capacity that can be developed and improved has been termed *incremental theory.*

Research consistently shows that those with an entity theory view respond to achievement situations in very different ways from those with incremental theory views. In general, entity theory tends to lead to suboptimal outcomes, whereas incremental theory is associated with more adaptive functioning. For example, viewing ability as a fixed capacity leads individuals to focus on proving that they possess the ability in question, and to worry that failure or low performance means they have (and will always have) low ability. In contrast, those who see ability as something that increases as one learns and develops skills and competencies tend to be focused on learning and increasing their competence, and they tend to view failure or low performance as a challenge rather than an indication that they lack ability. It is not surprising, therefore, that those with an entity theory of ability experience more negative emotion, lower self-esteem, and less effective coping strategies when they experience difficulty than do those with an incremental theory.

Whereas early research focused primarily on implicit theories about intelligence, subsequent work indicated that individuals often hold different assumptions in different domains. For example, one might have an entity theory with regard to mathematics and an incremental theory with regard to athletics. Therefore, contemporary scholars tend to focus on implicit theories with respect to specific abilities. Implicit theories with regard to any particular domain exert a substantial influence on how individuals construe achievement in that domain. In this sense, they are basic assumptions that affect the reasons individuals persist, or give up, the outcomes

on which they focus and the meanings they attach to those outcomes.

In sum, the aforementioned constructs are all centrally important in the literature on competence motivation. Constructs such as achievement attributions and evaluation anxiety certainly warrant attention, but they were omitted because they are best conceptualized as situation-specific processes rather than structural elements of competence motivation per se. Likewise, contextual factors (e.g., characteristics of the achievement environment) and cultural factors (e.g., culture-based sources of the meaning of competence) deserve consideration but were omitted because they focus on external, rather than internal, factors affecting competence motivation.

Integrative Models

Much of the existing research on competence motivation uses a single construct, linking it to processes and outcomes in achievement settings. However, some theorists have looked to integrate two or more constructs into an overarching conceptual framework. One such model will be briefly overviewed herein, specifically, the hierarchical model of approach-avoidance achievement motivation.

Achievement goals are the centerpiece of the hierarchical model because these goals are viewed as the proximal predictors of achievement outcomes. Achievement goals are viewed as concrete, situation-specific variables that account for the direction of competence-based pursuits. However, these variables cannot satisfactorily account for achievement behavior alone; other variables are needed to address the energization question—why people commit to competence in the first place. Higher-order variables such as intelligence/ability, self-perceptions of competence, competence motives, and implicit theories of ability are useful for this purpose. These higher-order variables, along with task-specific competence perceptions and competence values, lead individuals to adopt specific types of goals in achievement situations. These higher-order and task-specific variables are not thought to have a direct impact on achievement outcomes, but are thought to have an indirect influence through their effect on achievement goal adoption. These competence goals *are* expected to have a direct influence on achievement outcomes, as mediated by anxiety, perceived performance, achievement attributions, and other processes. Actual and

perceived performance are expected to moderate the influence of achievement goals on reactions to competence feedback, as well as the influence of initial achievement goals on subsequent achievement goal adoption.

An example will help illustrate this hierarchical regulation of achievement behavior. Low intelligence (from a relative standpoint), low self-perceived academic competence, fear of failure, and an entity theory of ability represent semi-independent reasons for why an individual might adopt a performance-avoidance goal for an upcoming achievement task. Low task-specific competence perceptions and high utility value might also contribute to performance-avoidance goal adoption. Once adopted, the performance-avoidance goal sets a framework for viewing the task in terms of the presence or absence of a negative normative outcome, and it undermines performance attainment by evoking anxiety that distracts the individual from optimal task engagement. Upon receiving negative performance feedback, those with performance-avoidance goals experience shame and strengthen their intention to avoid doing worse than others on subsequent tasks.

Models of competence motivation are of theoretical importance because they help to explain and predict competence-relevant behavior in a manner that is systematic and that generates new research and knowledge. These models are also of practical importance, because they highlight the fact that factors besides intelligence/ability have an important impact on achievement outcomes.

Andrew J. Elliot and Ista Zahn

See also Anxiety; Competition; Failure, Effects of; Goals

Further Readings

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: W. H. Freeman.

Covington, M. V. (1992). *Making the grade: A self-worth perspective on motivation and school reform.* Cambridge, UK: Cambridge University Press.

Dweck, C. S. (1999). *Self theories.* Philadelphia: Psychology Press.

Eccles, J. S. (1992). School and family effects on the ontogeny of children's interests, self-perceptions, and activity choices. *Nebraska Symposium on Motivation, 40,* 145–208.

Elliot, A. J., & Dweck, C. S. (Eds.). (2005). *Handbook of competence and motivation.* New York: Guilford.

Heckhausen, H., Schmalt, H.-D., & Schneider, K. (1985). *Achievement motivation in perspective* (M. Woodruff & R. Wicklund, Trans.). New York: Academic Press.

Lewin, K., Dembo, T., Festinger, L., & Sears, P. S. (1944). Level of aspiration. In J. M. Hunt (Ed.), *Personality and the behavior disorders* (Vol. 1, pp. 333–378). New York: Ronald Press.

McClelland, D. C. (1987). *Human motivation.* New York: Cambridge University Press.

Nicholls, J. G. (1989). *The competitive ethos and democratic education.* Cambridge, MA: Harvard University Press.

Motivation and Emotion

Using a variety of theoretical perspectives, this entry provides an overview of the interrelationships of motivation(s) and emotion(s) within academic learning contexts. Although this overview is not exhaustive, the frameworks will provide a foundation for understanding interlinks among these constructs.

Goal Striving

As a starting point for understanding connections among students' motivations, emotions, and learning, it may be helpful to think about academic motivation within a larger context of striving toward life goals. Although less scholarly attention has focused upon academic emotions and motivation, many scholars have discussed the role of emotions within the context of general goal striving. Within this overarching context, scholars have described ways in which much of human behavior is based on the attainment of one's goal aspirations and the emotions related to those endeavors. In this sense, goal striving motivates behavior; goal attainment brings satisfaction (i.e., pleasant/positive emotions), and failure in goal attainment brings dissatisfaction (i.e., unpleasant/negative emotions). But emotions are connected to more than feelings of satisfaction and dissatisfaction related to obtaining goals. Robert Emmons has suggested that all aspects of goal-directed behaviors have emotional links. Emotions are interlinked with one's commitment to goals, are involved with one's motivation during goal-related activities, and may serve as feedback informing one of the status of his or her goal attainment.

Although these constructs were first discussed in relation to general goal striving, one can see how they can be applied to learning-related goals as well.

Bernard Weiner's attribution theory was one of the first to focus on relationships of specific emotions with specific motivational behaviors in students' academic learning.

Attribution Theory

Weiner proposed that students' emotional reactions and motivational strivings are derived from evaluations that they use to explain a successful or failed attempt at a task or goal (i.e., causal attributions). Students' retrospective judgments of the causes of a successful or failed outcome affect their emotional responses and subsequent motivations in systematic and predictable ways. According to Weiner's theory, individuals make attributions along three dimensions: locus of control (internal–external), stability (stable–unstable), and controllability (controllable–uncontrollable). Additionally, he proposed that specific attributional patterns create specific, consequent emotions and motivations. For example, if a student attributes her successful task completion to her intelligence (internal and stable attributions) and because she put forth effort (controllable attribution), then she will likely experience pleasant emotions (i.e., happiness, pride) about her success and attempt similar, more difficult tasks in the future. On the other hand, if she perceives that she succeeded because the task was easy (external, unstable, and uncontrollable attributions), she may experience ambivalent emotions (e.g., satisfaction, but not pride). In the future, she may not attempt a similar, but more difficult task, because she may fear failure due to uncertainty regarding her abilities (internal and stable attributes).

Research has demonstrated that individuals' successes on challenging tasks leads to especially positive feelings, whereas their failures on easy tasks lead to especially negative feelings. In essence, attaining a difficult goal provides higher levels of pleasant/positive emotions such as satisfaction, joy in accomplishment, and feeling proud of one's self. On the other hand, when individuals fail to achieve easy tasks, they may generalize the implications of the failure, leading them to question their general competence and experience unpleasant/negative emotions such as worry, fear, and shame. For example, if a student perceives that a failure is due to internal, uncontrollable, and stable causes ("I'm not smart enough—I just don't have what it takes"), he or she may feel shame and will have no reason to be motivated to attempt the task again

(i.e., there is nothing that one could do to remedy the causal factors). Accordingly, the individual may respond with "helpless" behavior and have no further motivation for attempting a similar task in the future. On the other hand, internal, unstable, and controllable attributions for a failed attempt ("I didn't try hard enough") are more likely to induce feelings of guilt that do not call into question one's ability. Therefore, the student will more likely attempt similar, future challenges, perhaps with increased motivation.

Weiner's attribution theory has received empirical support, but the evidence has been inconsistent. One reason for the inconsistency of research support is that individuals have differing beliefs about the extent to which a particular attribution has stability or controllability. For example, attribution theory depicts intellectual ability as an internal, stable, and uncontrollable attribute; however, research has demonstrated that some students believe that intellectual ability is an acquirable skill that is developed through effort. For these students, attributions for failure that focus on their ability would not be interpreted as stable and uncontrollable, but would be seen as similar to effort, unstable and controllable. In this case, when these students experience failure, their attributions of low ability may cause them to experience disappointment, but not shame. Therefore, their attributions of low ability may cause them to increase their efforts or try different strategies, thus maintaining or increasing their motivation for future, similar tasks.

Although the full model of attribution theory has received variable support, most education theorists and practitioners agree that when students make attributions for their successes and failures that focus on internal and external control (i.e., place responsibility for the outcome on themselves or others) as well as attributions that focus on their ability and effort, their attributional processes may affect their emotions, motivations, and, ultimately, their academic learning and academic achievement.

A more expanded view of ways that students' emotions interrelate with their motivations and learning has been described through expectancy-value theories, explained in the next section.

Expectancy-Value Theories

John Atkinson, a forerunner in motivation theory, placed a great deal of importance on the impact of

anticipated pride and shame with respect to motivating behavior. In his theory of achievement motivation, he proposed that the *value* placed on successful attainment of a goal, linked with the strength of *expected* outcomes and emotions—either pride about anticipated success or shame about anticipated failure—are the driving forces (i.e., expectancy value) that determine the extent to which an individual will approach or avoid a task.

Following Atkinson, expectancy-value theories expanded the motivational variables that were considered within the traditional components of expectancy and value to include interactions of cognitions, motivations, and emotions that influence students' academic motivation. Current expectancy-value theories of academic motivation propose that there are three major motivational components of students' learning: an *expectancy* component (which includes students' beliefs about their capabilities and their expectations for success, e.g., self-efficacy); a *value* component (which includes students' goals, reasons, and interests regarding academic endeavors, e.g., mastery and/or performance goals, future and/or instrumental goals); and an *affective* component (which includes students' feelings regarding academic endeavors). The extent to which a student will enact motivated behavior will depend on the strength of a variety of expectancy-related and value-related components.

Extant research regarding self-efficacy (i.e., expectancy component) for learning has indicated that students' self-efficacy beliefs influence their acquisition of skills as well as their persistence at difficult or uninteresting tasks. Students who believe that they are competent and capable will demonstrate motivated behaviors even after they have experienced prior difficulties or received feedback indicating they are not doing well. Self-efficacy has been shown to influence students' choices about whether or not to engage in similar, future activities. In addition, research has shown that self-efficacy beliefs are integrally related to students' capabilities for self-regulated learning as well as their use of cognitive learning strategies and motivational (volitional) strategies.

Regarding the types of goals that students may have (the value component), research has shown that when students have internally rewarding goals such as wanting to master educational tasks, they engage in academic activity because the activity itself, or valuing the mastery, is motivating to them. For students who are motivated by primarily external goals (such as grades or praise), these rewards often provide external confirmation of their ability (or lack thereof). In contrast to students who are motivated by intrinsically rewarding goals, students whose academic motivation centers on extrinsic rewards tend to have less persistence when faced with challenges, tend to use more surface cognitive strategies (e.g., simple rehearsal for memorization with respect to test preparation), and also evince more worry. However, although students can be primarily intrinsically or extrinsically oriented in their academic goals, combinations of these orientations are often exhibited. In the current system of education, many activities that are not enjoyable within themselves may be viewed as being instrumental to reaching intrinsically rewarding, long-term, future goals (e.g., completing a high school degree, then applying to college, and completing all college requirements in order to fulfill a future employment goal). When a student fails a task and feels disappointment, guilt, or shame, he or she is more likely to be resilient in his or her motivation (e.g., increasing or maintaining motivation) if the task the student failed is instrumentally connected to an important future goal that has intrinsic incentives.

With respect to the affective component of expectancy-value motivation theories, Reinhard Pekrun has renewed an emphasis on the importance of emotions in achievement motivation. In his expectancy-value theory (later renamed *control-value theory*, recognizing that control-related issues may be connected to past attributions, current appraisals, or future expectancies), Pekrun proposed that emotions act as cognitive and motivational mediators within academic learning contexts. His theory explains how three categories of emotions, based on the time frames during which emotions occur, can affect motivation and learning:

1. Process-related emotions that happen during a learning task (e.g., enjoyment, boredom);

2. Prospective emotions that occur with anticipation to future outcomes (e.g., hope, anxiety); and

3. Retrospective emotions that occur after task completion (e.g., pride, shame).

In particular, control-value theory proposes that, no matter the time frame of the emotions, students' control-related and value-related appraisals are the main sources of students' academic emotions and subsequent motivations.

Perhaps one of the most important aspects of control-value theory is that it describes the reciprocal processes that occur among students' cognitions, emotions, and motivations that work together (or against each other) to influence students' ongoing learning processes. For example, students' cognitive appraisals of control and value influence their experiences of emotions and motivations. In turn, students' emotions affect their motivations and their cognitive processes for learning. Not surprisingly, control-value theory proposes that students' experiences of positive (i.e., pleasant) emotions lead to positive (energized), motivated behavior and successful academic achievement, which in turn lead to more positive emotions. In contrast, the theory proposes that the impact of negative (i.e., unpleasant) emotions is not necessarily a straight or simple path. This is because, in contrast to positive emotional outcomes that signal all is well, negative emotional outcomes invoke complex thinking and problem solving to discern what went wrong. The consequences of the complex thinking instantiated by negative emotions may lead to either motivated behavior or helpless behavior, depending on an array of motivation-related influences (e.g., self-efficacy, important future goals) or circumstances (e.g., situational constraints).

Moods and Cognition

Although less is known about specific educational emotion-cognition linkages, a body of evidence does exist with respect to general affective influences on cognitive processes. For example, research on mood induction has demonstrated that general positive and negative affect may affect information processing. Within this framework, researchers have proposed that information may be stored in memory according to its affective tone and then linked through affective nodes within neural networks. In support of this theory, research has demonstrated that individuals in negative mood states seemed to be "primed" to recall negative experiences, and they sought out information that helped them to maintain their "negative" moods. For example, not only did negative-mood-experiencing individuals spend more time reading negative information, they spent more time rereading negative details and remembering negative events. The opposite was true of individuals in positive moods; these individuals focused on positive details and remembered more positive events. Mood states have

been shown to have a strong influence on decision-making processes and self-perceptions of competence.

Mood states have also been shown to have an effect on attributional processes. For example, research has found that individuals in positive moods tended to attribute a successful past performance to internal attributions and unsuccessful past performances to external attributions. Additionally, individuals in negative moods tended to make internal attributions for their own past failures, but not the past failures of others.

In addition to research conducted with respect to mood-congruent information processing and outcomes, research has been conducted with respect to the influence of moods on type and depth of information processing. In a variety of situations, research has demonstrated that individuals in an elated or positive mood were less likely to use systematic, detail-oriented information-processing strategies, whereas people in mildly depressed or negative moods were more likely to use systematic, detail-oriented information-processing strategies. For example, in contrast to individuals in positive mood states, individuals in negative mood states used more elaborative information processing in response to strong persuasive messages, used detailed information, and were more accurate on a performance appraisal task with respect to perceptions and appraisals of others. Additionally, individuals who were in negative mood states were more likely to calculate the costs and benefits of helping (i.e., used more cognitive processes) when deciding if they should be helpful.

Emotion and Cognitive Capacity

One way in which emotions may affect students' learning is by compromising cognitive availability. In other words, cognitive capacity may be involved with the direction and maintenance of attention in academic situations. In line with theoretical propositions of the impact that emotions have on learning, cognitive capacity theory states that cognitive components of emotions (e.g., appraisals or self-efficacy) can particularly occupy a person's mental resources and thereby consume cognitive availability. The idea that emotional processing may encumber cognitive capacity is supported by Kahneman's Cognitive Capacity Model of attention as well as Kanfer and Ackerman's model of self-regulation.

The Cognitive Capacity Model proposes that an individual's capacity to think is limited. This theory

recognizes that different mental activities demand different requirements of cognitive load for processing. Additionally, the arousal system is considered to be connected to the cognitive system; therefore, arousal and cognitive capacity fluctuate according to the changing demands of cognition and arousal. In this view, certain emotions may cause an increase in the arousal system (e.g., high anxiety), occupying cognitive capacity that may be required for academic activities. In light of this model, high-intensity emotional processing could be viewed as consuming cognitive resources by taking away from on-task attentional resources, allocating resources to off-task activities (emoting), or interfering with resources for cognitive self-regulatory strategies that are needed for academic activities. If students cannot regulate their affective arousal, their attentional capacity to conduct academic tasks may be reduced.

Future Directions

The interrelationships of students' emotions, motivations, and learning are indeed complex. The dynamic complexity of these constructs has been shown to emerge among theoretical frameworks from the fields of goal-striving theories, attribution and appraisal theories, academic motivation theories, and cognitive processing theories. Research focusing on understanding the complexity of cognitive, motivational, and emotional forces within students' academic learning is currently in progress and is poised to add new insights into the dynamics involved in individuals' learning-related processes.

Jeannine E. Turner and Joel B. Goodin

See also Emotional Development; Emotional Intelligence; Learning; Maturation

Further Readings

Atkinson, J. W. (1964). *An introduction to motivation.* Princeton, NJ: Van Nostrand.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice Hall.

Bower, G. H. (1981). Mood and memory. *American Psychologist, 36*, 129–148.

Deci, E. L., & Ryan, R. M. (1991). A motivational approach to self: Integration in personality. *Nebraska symposium on motivation* (Vol. 38, pp. 237–288). Lincoln: University of Nebraska Press.

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review, 95*, 256–273.

Kanfer, R., & Ackerman, P. I. (1989). Motivation and cognitive abilities: An integrative/aptitude-treatment interaction approach to skill acquisition. *Journal of Applied Psychology Monograph, 74*, 657–690.

Op't Eynde, P., & Turner, J. E. (2006). Focusing on the complexity of emotion issues in academic learning: A dynamical component systems approach. *Educational Psychology Review, 18*(4), 361–376.

Schwarz, N., & Bohner, G. (1996). Feelings and their motivational implications: Moods and the action sequence. In P. M. Gollwitzer & J. A. Bragh (Eds.), *The psychology of action: Linking cognition and motivation to behavior* (pp. 7–26). New York: Guilford.

Turner, J. E., & Schallert, D. L. (2001). Expectancy-value relationships of shame reactions and shame resiliency. *Journal of Educational Psychology, 93*(2), 320–329.

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review, 92*, 548–573.

MOTOR DEVELOPMENT

Over the first years of life, humans exhibit an increasing ability to coordinate movements of the body. The initial emergence of these motor capacities and their subsequent refinement is often referred to as motor development. The physical development of muscle strength is required for this, but visual perception and skills such as postural control and balance are critical as well. The dynamic systems approach to development has been extremely influential in this area, as it has emphasized the need to consider interactions between physical, perceptual, cognitive, social, and motor development in order to understand any of the individual domains.

In adulthood, motor performance peaks and then gradually declines; motor development thus occurs across the life span. The vast majority of research on motor development, however, has focused on the early increases in motor capacity rather than its later deterioration. This entry does so as well.

Over the past several decades, an increasing amount of research has been devoted to characterizing and understanding human motor development. A good description of the ages at which different motor capacities emerge has been compiled, at least within the context of North American and Western European

culture. Within this cultural context, the order and approximate ages at which different behaviors first emerge is quite consistent (see Table 1). This consistency has led to claims that motor development is largely mediated by genetically specified maturation, but cross-cultural and experience enrichment studies have both demonstrated the important roles that cultural convention and environmental support play in motor development.

For instance, in the United States, it is very common for infants to begin crawling between 5 and 7 months of age. In an infant population sampled in urban China, however, the onset of self-produced locomotion was found to occur approximately 3 months later. This is almost certainly not due to genetic differences between American and Chinese infants, but rather to an urban Chinese cultural tendency to set infants down only on thickly cushioned surfaces, on which crawling is more difficult. In other cultures, for instance the Kipsigis of Kenya, babies hold their heads up and walk earlier than North American children. Kipsigi parents deliberately teach these skills to their infants, for instance seating them in holes dug in the ground and surrounding them with blankets to hold them upright. By frequently bouncing children on their feet, Kipsigi parents also promote earlier walking behaviors. As with many other aspects of development, it seems that the course of motor development is influenced by interactions between genetic and environmental influences (i.e., nature and nurture). Researchers now believe that motor development is the outcome of a complex interplay between central nervous system processes, physical and physiological variables, and relevant environmental circumstances.

The progression of motor development has often been described as consisting of four phases. (1) Infant development progresses from reflexes to rudimentary movements such as sitting, crawling, creeping, standing, and walking. (2) Fundamental skills such as running, climbing, jumping, balancing, catching, and throwing are next to develop. (3) During late childhood, more specific movement skills will appear, and the general fundamental skills become more refined and appear more fluid and automatic. (4) In adolescence, the specific movements develop further and become more specialized. Specialization depends on talent, motivation, and the amount of practice with a specific skill.

Within all four phases, motor development is often further divided into categories of gross and fine

Table 1   Approximate Ages of Early Motor Skill Onset

Approximate Age (months)	Skill
0–1	Control eyes to look at targets of interest; newborn reflexes
1–3	Lift up head; grasp objects placed in hand; use eyes to smoothly track moving targets; lift self using arms from a prone position; roll from side to back
3–5	Roll over in one direction; sit up if propped in place; grasp objects
5–8	Sit without support; stand while grasping support; one-handed grasping
5–10	Pull self to stand; point at objects; grasps with thumb and finger
5–11	Crawling; direct food to mouth; use eating utensils; build tower of two cubes
10–14	Stand without support; walk; build with blocks; put objects into containers; pick up very small objects (e.g., raisins, blades of grass)
13–18	Walk backwards and sideways; run; climb; walk up stairs; kick a ball while walking; scribble with crayons on paper
18–30	Jump; skip; walk on tiptoe; pick up small objects

Note: Within North American and Western European cultural contexts, the order of motor skill onset is quite consistent, as are the ages at which children first exhibit these behaviors.

motor skills. Gross motor skills include behaviors that involve movements of large portions of the body, such as lifting one's head, rolling over, sitting up, balancing, crawling, and walking. Fine motor skills include behaviors that emphasize hand-eye coordination, such as manipulating small objects, transferring objects from hand to hand, drawing, cutting, and threading beads. Gross motor performance typically precedes fine motor skill, perhaps because gross motor skills must be performed in order to support most fine motor tasks. For example, one must be able

to effectively sit up in order to have both hands free for manipulating objects.

In the following sections, eight major topics in motor development are considered. This list certainly does not provide a comprehensive account of motor development research, but is intended to provide a sense of the range of issues that have been and continue to be explored in this area. The first section considers the foundations of motor performance inherent during prenatal and early postnatal development. Subsequent sections describe the onset of crawling and walking behaviors and consider how these and other motor developments interact with cognitive, social, and emotional development. Then the emergence of the fine motor skills of drawing and writing is considered, along with a summary of motor development associated with late childhood and adolescence.

Prenatal Motor Development

The human embryo typically begins to move independently in the second postconception month. The embryo is too small at this stage for the mother to feel the movements, but they are often made in response to sensory stimuli such as sounds or pressure. Even at this early stage of development, the link between sensory and motor systems is evident. In the third month of prenatal development, the movements become increasingly frequent and coordinated. Infants will kick, form a fist, and even engage in thumb sucking. By the fifth prenatal month, a human fetus will hold up its hands to shield the eyes in response to a bright light, showing that this visuomotor reflex is already in place.

It is commonly presumed that prenatal motor development must be a purely genetically specified maturational process, because the womb shields the baby from the external environment that might otherwise influence development. However, there are several examples that suggest this is not true, at least one of which can be observed in the motor domain of "handedness." Between 1 and 2 years of age, many children will begin to exhibit a consistent hand preference; most children will be right-hand dominant. The preference does not become stable until this late age, but it can be predicted based on the prenatal position of the child in the womb. A fetus whose left side consistently faces out through the belly of the mother during the pregnancy is likely to be right-handed; whereas right-side-out infants will likely be left-handed. Several theories have been proposed to

explain this difference. For instance, the left and right ear auditory systems project neural activation to the left and right cortical hemispheres respectively. It is presumed that the ear that faces out receives more stimulation, resulting in greater development of the corresponding hemisphere. If the left cortical hemisphere becomes dominant, then the hand that this hemisphere controls (i.e., the right) will be dominant; or vice versa. Regardless of whether this theory is correct, it is clear that, even in the womb, environment plays a role in mediating motor development.

Overall, prenatal development of the motor system enables the infant to perform certain actions even prior to being born. The extent to which genetic and environmental influences play a role in this development, however, is still not fully known.

Newborn Motor Capacities

Infants are born with a set of basic motor abilities, but they are primarily activated by specific external stimuli. That is, they are reflexive responses rather than voluntary actions. For instance, if one touches an infant's cheek near the mouth, she will turn her head in that direction (the "rooting reflex"). If the infant's lips are stimulated by physical contact, she will engage in a sucking behavior (the "sucking reflex"). This combination of reflexive responses, like many others, is believed to have adaptive value and is often interpreted within the context of evolutionary theory. For example, the two reflexes described here, rooting and sucking, will help a child to gain nutrition in the context of feeding at the breast. In one fascinating study, researchers found that when a newborn was placed on the mother's stomach and allowed enough time, the combined effect of several inborn reflexes caused the infant to move forward, latch on to the breast, and begin to nurse—a feat termed *self-attachment.*

The infant reflexes that are predominantly used for protection, nutrition, or survival are called the *primitive reflexes.* These reflexes are usually present at birth, but most vanish after a few weeks or months. The disappearance is often attributed to the maturation and increased activity of the cerebral cortex, which inhibits reflexive responses. Pediatricians commonly use these reflexes as diagnostic tools for assessing the infant's level of neurological development. Severe deviations from the normal time line of reflex onset and disappearance may indicate neurological immaturity or dysfunction.

Postural reflexes do not provide the immediate adaptive value of primitive reflexes but support the development of later voluntary movement. For example, the "stepping reflex" can be observed in most full-term babies. When the infant is held so that the feet are flat on a surface, the infant will lift one foot after another in a stepping motion. Many developments must occur over the first year of life to make independent walking possible, but newborn stepping behaviors demonstrate that the brain systems and motor patterns that will produce that walking are already present at birth.

Like the primitive reflexes, stepping seems to disappear at around 2 months of age. Many researchers have attributed this to cortical maturation, but several clever studies have called this interpretation into question, both for the stepping reflex and for others as well. When a 2-month-old, who does not normally display the stepping reflex, is submerged up to the chest in warm water, the stepping reflex returns. This finding suggests that the disappearance of the stepping reflex may not be due to cortical maturation, but instead to the simple physical development of heavier legs. When the infants' legs are submerged in water, the muscular strength required to lift the legs is greatly reduced, and the reflexive behavior returns. This type of experiment greatly emphasizes the value of a dynamic systems approach to motor development, which considers how motor performance is influenced by the interaction of many different types of developmental changes.

Other newborn motor capacities suggest the presence of impressively complex visuomotor processing. For instance, one study found strong evidence that newborns imitate common facial expressions, such as smiling, mouth opening, and tongue protrusion. That is, when a newborn views an adult sticking out her tongue, the probability that the infant will also perform this facial expression is greatly increased. Such a feat is remarkable because the newborns who participated in these studies had seen only a few faces prior to the start of the procedure, they had never seen their own faces, and they had certainly not had the opportunity to learn associations between certain muscle contractions and changes in their own facial appearance. It seems that all of the components necessary for facial imitation are specified in the genome of the newborn. Just how a facial imitation system might be encoded at the level of DNA is essentially unknown.

Infant Reaching

One of the most salient motor development milestones of the first 4 months is the emergence of visually guided grasping. Even during the first weeks of life, children exhibit prereaching behaviors, such as reaching out to swipe at a target of interest. Not until approximately 4 months, however, do they reliably reach out and grasp a target such as a small toy. Children make their first successful grabs using both hands, but soon after they will begin grasping targets using only one hand. When they do, they typically use "ulnar" grasps, relatively clumsy movements in which all of the fingers are pressed against the palms. Within a few months, ulnar grasps will be replaced by "pincer" grasps that make use of the tips of the thumb and fingers.

The cascade of reaching development shows an increasing adaptation of reaching to the physical constraints of the physical environment. For instance, when an adult reaches for a tall, thin target, she typically rotates her wrist so that the opening of the grasp corresponds to the orientation of the target. Children's grasps do not exhibit this characteristic at 5 months of age, but do so by 9 months of age. The grip apertures of adult grasps are precisely correlated with the sizes of reaching targets. Infants' grip apertures are not highly correlated with target size at 7 months of age, but this tendency emerges by 9 months.

Not all aspects of reaching control develop so long after the onset of grasping, however. Perhaps the best example of this is the development of the ability to catch moving targets. If one wishes to catch a moving target, a reach cannot be aimed at its current location. (By the time the reach is completed, the target will have moved on to a different location.) A reach for a moving object instead must be aimed ahead of the target, along its future path of motion such that the paths of the hand and the target intersect the same position at the same time. Given the complexity of this, one might presume that catching moving targets would be more difficult than grasping stationary objects. As soon as infants are able to grasp stationary objects, however, they are immediately able to catch moving targets as well. Indeed, some studies of prereaching behaviors suggest that this future-oriented aiming of motions is in place before successful grasping emerges.

Crawling and Walking

Between 5 and 7 months of age, most infants begin to exhibit a variety of crawling behaviors (belly

crawling, hands-and-knees crawling, directed rolling, etc.). As they become independent movers, their relationship to the world around them changes in several fundamental ways. With this in mind, it is perhaps sensible that a host of other behavioral changes accompanies this development. For example, post-crawling infants begin to respond to pointing actions in an adult-like fashion, looking at a target to which an experimenter points rather than at the experimenter himself. Crawling children also make more use of visual information for maintaining posture than do noncrawlers. Crawlers are better at retrieving hidden objects and more able to focus attention on distant targets, and they exhibit a wider range of emotional expressions.

One related change that has been studied extensively is the infant tendency to avoid a "visual cliff." Novice crawlers will crawl over the edge of a drop-off without much hesitation. Only after some weeks of crawling experience does a healthy fear of heights emerge. Interestingly, when children begin to walk a few months later, they need to relearn this cliff avoidance. The motor learning that has taken place in the crawling domain does not transfer to the task of walking, suggesting that human motor performance may be mediated by somewhat "modular" mental systems.

Interactions Among Cognitive, Social, Emotional, and Motor Development

Infant motor performance is clearly subject to the principles of operant conditioning. This has been found in children as young as 3 months of age, who will engage in a variety of behaviors more often (e.g., kicking, arm waving, head turning) if the behavior is linked to (i.e., contingent with) an interesting sensory event. For example, if an infant's ankle is attached to a mobile with a ribbon, such that kicking the leg causes the mobile to shake, the infant will learn this association in a matter of minutes and engage in more frequent kicking. Connections between learning and motor performance such as this are interesting for at least two reasons. First, motor performance provides a tool for assessing the learning abilities of preverbal children. Second, it makes clear the ongoing refinement of motor actions that takes place over the course of development. Whenever a child is engaging in motor behaviors, even in a play capacity, the outcomes

of the behaviors serve as a tool for refining later performance. Jean Piaget's extensive and influential theories of human cognitive development progressed from this foundation. Piaget presumed that all knowledge starts from a basis of sensorimotor experiences and learning with the world.

Researchers' knowledge of children's earliest ability to recall prior events (as opposed to simply recognizing them) comes entirely from studies of the motor task of "deferred imitation." After watching an adult perform a specific novel action, infants are more likely to perform the same action later when presented with the same stimulus materials. For example, if a 6-month-old watches an experimenter shake a toy such that it makes an interesting sound, then when the infant is later presented with the same toy, even after a 24-hour delay, she is more likely to shake it in the same way. This is commonly considered to be the earliest indicator of "declarative" memory. Interestingly, 6-month-olds' memory performance in this task seems remarkably context-specific. That is, if the same stimulus materials are presented in a different room, then the likelihood of deferred imitation is essentially eliminated. Adults also show better recall of information in the same context in which it was learned, but the size of this effect is relatively modest. By 9 months of age, the context-specificity of deferred imitation is greatly reduced, suggesting that the mental representations that drive deferred motor imitation develop substantially during this period.

As children eventually develop the ability to plan and reason about the world, the best indicator of their level of cognitive performance is often not their verbal behavior, but their general motor responses. Over the first several years of life, children are remarkably adept at language learning, seeming to absorb the vocabulary, syntax, and pragmatics necessary to communicate with only a modicum of explicit effort. Children usually express this learning with increased verbal utterances, but some children are also taught a sign language (e.g., American Sign Language or ASL). It is notable that the level of linguistic performance, in terms of both production vocabulary and level of grammatical complexity, is consistently better in the sign language domain than in the verbal domain. Producing the sounds of human language is a complex and rapid process, mediated by the use of dozens of different muscle groups to produce sound and "filter" it via rapid movements of the throat,

tongue, and mouth. A limiting factor of linguistic performance thus seems to be the ability to control the rapid and complex muscle contractions required to verbally communicate.

Emotional development is also better understood within the context of motor development. As children develop new action abilities, they provide conduits to new types of social and emotional interactions. For instance, when children learn to reach for objects, they will often begin to have their first experiences with parents scolding them for grabbing things that they should not. As children learn to crawl, they become more autonomous and goal directed, creating even more situations in which they may interact and sometimes conflict with the wishes of their parents. It is thus not surprising that the earliest onset of behaviors associated with anger have been found to be correlated with crawling.

Motor development does not occur in a vacuum. Motor behaviors take place in a social and emotional context that can influence that development. Similarly, motor development has often been found to be a catalyst for that social and emotional development.

Correlations Between Brain and Motor Development

The study of brain development has progressed rapidly in recent years, as new techniques for exploring adult and child brain activity have become more accessible. Studies of brain activation in adults have identified that the cerebellum and basal ganglia are areas critical to motor learning performance. In children, the level of activity in these brain regions is correlated with the level of behavioral proficiency.

The corpus callosum is a large brain structure that connects the left and right cortical hemispheres, enabling information to pass from one to the other. As this structure develops, children's ability to perform actions in which the left and right hands are coordinated greatly improves.

Several studies have identified other changes in brain development that coincide with, or even precede, changes in behavior. For instance, neuronal connections within the frontal and occipital lobes of the brain are rapidly produced around the ages of 4 months, 7 months, and 12 months—the ages of onset for reaching, crawling, and walking. These bursts of cortical activity appear to begin in the frontal lobes and then spread to other areas of the brain.

Drawing and Writing

By age 2, many children start to draw on paper using crayons or other implements, usually in the form of arbitrary scribbling. By age 3, children will begin to produce representational forms, although the initial forms are often produced by accident. For instance, a child engaging in scribbling may notice that she has inadvertently produced a shape that she can identify by name, such as a circle or an oval. By age 4, children commonly produce such representational forms intentionally. Four-year-olds' drawings are almost always quite stylized and inaccurate, but are typically recognizable to naïve observers. For example, children commonly draw people as circular heads with small arms and legs, but without defined bodies.

The level of detail and accuracy of children's drawings gradually increase over subsequent years. By age 5, most children can copy two-dimensional shapes and incorporate them into their pictures. Some 5-year-olds show a rudimentary use of three-dimensional depth information, such as drawing more distant objects as smaller than near objects. Not until age 9 or 10 years, however, do children realistically depict three-dimensional space. Only at this age do children draw nearer objects as occluding more distant objects; some linear perspective cues will also be present in these later drawings.

During the school years, children can also be taught to write numbers and letters. In the early stages of writing, children typically only write very large letters and numbers, due to their tendency to control writing using movements of the whole arm. As they develop the ability to write using primarily movements of the wrist and fingers, the print size will shrink down to a more adult-like size.

All of these fine motor capacities are clearly influenced by schooling, as well as the level of emphasis placed on writing and drawing by a child's parents as well as the society in which he or she lives. These parental and societal influences are also apparent for many aspects of motor development in late childhood and adolescence.

Motor Development in Late Childhood and Adolescence

Through late childhood and adolescence, the bones lengthen and undergo additional calcification, producing a more rigid frame for the skeletal muscles; the

muscles grow stronger as well, enabling the child to produce greater amounts of force. The shape of the body also changes, as the arms and legs lengthen relative to the size of the torso. All of these factors contribute to a general increase in balance, flexibility, agility, and maximum force production across this age range.

Although there are clear physical changes across this age range, few milestones of motor development have been identified beyond the age of 8 years, and with good reason. There are certainly differences between the level of motor coordination found in adults and children, but most of these differences are quantitative rather than qualitative in nature. In the sports of gymnastics and diving, two activities that demand extremely high levels of motor performance, it is not uncommon to find world-class performances among children as young as 13 years of age. The primary limits on motor performance beyond middle childhood thus seem to be amount of talent and expertise. On average, older children will perform at a higher level than younger children, but this seems primarily due to the accrual of experience rather than a disappearance of physical or cognitive limitations.

Implications

Many of the earliest studies of developmental psychology were studies of motor development. Motor performance can be directly observed, and it is produced by children of all ages. Motor performance is a central component of all behaviors, and it provides a tool for exploring many aspects of cognitive and emotional development. The dynamic systems approach to the study of human behavior began in the domain of motor development, but a growing number of researchers have found it fruitful to consider non-motor aspects of development from this perspective as well. Motor development has thus played a central role in the historical progression of developmental psychology and seems likely to continue to be a pervasive aspect of developmental science in the coming decades.

Peter M. Vishton and Natalie H. Brito

See also Athletics; Brain-Relevant Education; Dynamical Systems; Maturation; Measurement of Cognitive Development; Myelination; Neuroscience; Perceptual Development; Physical Development

Further Readings

Adolph, K. E., & Berger, S. A. (2006). Motor development. In D. Kuhn & R. S. Siegler (Eds.), *Handbook of child psychology. Vol. 2: Cognition, perception, and language* (6th ed., pp. 161–213). New York: Wiley.

Campos, J. J., Anderson, D. I., Barbu-Roth, M. A., Hubbard, E. M., Hertenstein, M. J., & Witherington, D. (2000). Travel broadens the mind. *Infancy*, *1*, 149–219.

Fischer, K. W., & Bidell, T. R. (1998). Dynamic development of psychological structures in action and thought. In R. M. Lerner (Ed.), *Handbook of child psychology. Vol. 1: Theoretical models of human development* (5th ed., pp. 467–561). New York: Wiley.

Thelen, E., & Fisher, D. M. (1982). Newborn stepping: An explanation for a "disappearing reflex." *Developmental Psychology*, *18*, 760–775.

Vishton, P. M., Ware, E. A., & Badger, A. N. (2005). Different Gestalt processing for different actions? Comparing object-directed reaching and looking time measures. *Journal of Experimental Child Psychology*, *90*, 89–113.

Multicultural Classrooms

Historically, the idea of multiculturalism has been defined in many different ways. For some, the label refers specifically to ethnic group differences; for others, it spans such concepts as socioeconomic status, gender, sexual orientation, age, disability status, or any demographic data that differentiate one person from another. All of these are valid areas of multicultural study, but two in particular (economics and ethnicity) have the largest impact on educational performance. Thus, this entry addresses culture mostly in terms of ethnic and economic differences. It should also be noted that, when specific examples of ethnic differences are provided, only White, African American, and Hispanic people are discussed in any detail because these three ethnicities make up more than 80% of the general population of the United States. This is not to say that other aspects of multiculturalism and other ethnic groups do not have an impact that should be addressed, but the topics discussed here were chosen to be most representative of the experiences in the average multicultural classroom.

Just as the term *multiculturalism* has no single, specific definition, the term *classroom* has become more difficult to generalize into a single concept as well. Classrooms can be housed in traditional brick-and-mortar buildings, or they can be housed as space

on a hard drive in a virtual location created with online technology. They can be used to teach preschoolers or postdoctoral students. Classrooms may serve several hundred children at once or only one student at a time. Subjects may range from vocational and artistic pursuits to traditional academic studies. Although all of these are legitimate descriptors for a classroom, for the purposes of this entry, a classroom is defined as a physical room with one teacher and 20–30 children in a primary or secondary school building. These types of classrooms usually involve three parties—teachers, parents, and children—whose individual experiences will be described in greater detail.

Children's Ethnic Differences in the United States

What does a typical multicultural classroom look like in terms of ethnic and economic differences? According to data from the 2000 U.S. Census, from the general population approximately 25% of those surveyed described their race as non-White. In addition, those who described themselves as non-White showed a higher probability of living in poverty. For example, the poverty rate in 2005 was approximately 10%–15% among the general population; however, the numbers look conspicuously different when examined by cultural groups. Approximately 25% of African Americans and 20% of Hispanics live in poverty. The difference in poverty rates is more striking for the 19% of children living in the United States with at least one parent who is foreign-born. Approximately 33% of children who are native to the United States but born of parents who were born outside of the United States live in poverty; compare this number to approximately 40% of children who live in poverty who live in the United States but were born outside of the country. Given the strong ties between ethnicity and poverty, it is frequently difficult to discuss ethnicity differences without discussing socioeconomic differences as well.

Socioeconomic differences frequently lead to lifestyle differences; therefore, many ethnically diverse students live in a manner very different from the majority culture. For example, a White individual's annual income falls at approximately $24,000 compared to $14,000 for African Americans and $12,000 for Hispanics. In the general population, approximately 75% of children live in a household with their married parents. However, this number drops for specific ethnic groups

such as Hispanic children (approximately 65%) and African American children (approximately 38%). Health insurance is another major lifestyle difference. Approximately 7% of White children are uninsured, compared to 13% of African American children and 22% of Hispanic children. Parental levels of education change based on ethnicity as well. Approximately 36% of White adults have a bachelor's degree compared to only 19% of African American adults and 14% of Hispanic adults. These data are important because research consistently shows that parental education has a huge impact on student success and predicting graduation rates.

Compare the statistics listed above for multicultural children and parents to statistics for teachers. The National Center for Education Statistics states that the average teacher's annual income is approximately $43,000. This is a 350% increase in income compared to the average Hispanic parent. They add that only 16% of teachers describe themselves as non-White. More than 98% of teachers hold at least a bachelor's degree, and 46% hold a master's degree or higher. These educational and economic differences influence teachers' daily lives, view of the world, classroom expectations, life experiences, and feelings of safety and security. Given such significant differences between the typically White, middle-class teachers and their more culturally diverse students and parent collaborators, each individual's experiences in the multicultural classroom cannot help but differ widely.

Experience of Students in Multicultural Classrooms

The National Center for Education Statistics provides data on school achievement and school behaviors, broken down by race. These data show that ethnicity and school achievement are strongly related. For example, dropout rates were significantly higher for Hispanics (23.8%) than for either African Americans (11.8%) or Whites (10.3%). In addition, SAT scores were lower for African Americans (861 combined out of a possible 1,600) than for Hispanics (932 combined), and much lower than those for Whites (1,068 combined).

One theory for the cause of the discrepancies in educational achievement is that ethnically diverse children have different cultural learning styles that may or may not be compatible with traditional American

teaching methodologies. Guild writes about several specific cultural learning styles common to different ethnic groups. She cautions that these are generalities applied to groups, and that particular styles may not be relevant to every member of each group. Guild describes the notion that Mexican Americans tend to have higher regard for personal relationships and thus seek greater contact and interconnection with their teachers than other groups might. In addition, research suggests Mexican Americans learn best when information is presented in terms of generalities and patterns; rote memorization is more difficult for them.

Guild's descriptions of research on African American learning styles indicate that this group prefers physical activity and orally recalled information; therefore, African American students may do better in a class that emphasizes group activities and discussions over lectures. She compares both groups to White learners, who think best in terms of accuracy, individuality, and objectivity, and therefore are best suited for experiences such as testing and competition. Because the latter are the most common learning experiences within U.S. classrooms, this may partially explain why White students may perform better.

Another theory put forth by MacMillan and Reschly states that the differences in academic performance are not due to cultural learning-style differences as much as to economic ones. Supporting this theory are research findings that poverty brings with it a series of risks known to cause lowered scores on tests. These risks include medical differences such as poor prenatal care, low birth weight, drug and alcohol use, and increased physical abuse.

Also included are parent-related issues such as after-school availability, involvement in their children's school, necessary knowledge of academic material to help with homework, ability to provide outside learning experiences, knowledge of the English language, and use of complex vocabulary to teach higher-order language skills. Another parent-related issue is frequent relocation, which is common to any group living in poverty but may have the largest educational impact on families of migrant workers.

School-related issues have a large impact as well. Schools serving students of lower socioeconomic status tend to have fewer certified teachers, less experienced teachers, fewer class materials, and older facilities. These schools tend to have higher instances of violence, weapons, drug use, and teen pregnancy. These behavioral issues affect school performance in

that children worried about violence and children using drugs have a harder time paying attention to class materials and retaining knowledge.

Research supports the notion that the differences between school behaviors are strikingly different for multicultural students. The National Center for Education Statistics states that only 10% of White students report ever being engaged in a physical fight on school property, compared to approximately 17% of African Americans and 17% of Hispanics. Statistics also reveal a discrepancy in how safe children feel attending their own schools. Whereas only about 3% of White children felt too unsafe to attend school, 8% of African American students and 9% of Hispanic students described their schools as unsafe.

Payne describes some of these behavioral discrepancies as a mismatch between middle-class societal mores and those of individuals living in poverty. Specifically, she states that children living in poverty may speak to teachers in a way that seems inappropriate to the teachers. However, these students may become defensive if the teacher talks to them in a way that makes the students feel like they are being talked down to. In addition, students may spend more time socializing and making jokes than learning because humor and social skills are more highly valued in their community than is academic achievement. Finally, violence is described as a way of life that may be needed for protection, or it may just be more of a habitual response than an active desire to hurt someone else.

Successful Classrooms for Multicultural Students

Either poverty or cultural learning styles could be the primary cause of academic, social, and emotional differences among children in the multicultural classroom, but the true culprit is most likely a combination of all factors. Therefore, teachers of successful classrooms for multicultural students develop their instructional plans to work within a child's cultural learning style, available resources at home, personal strengths and abilities, and cultural knowledge base. This can be done through the use of differentiated instruction by which some students are taught differently from others within the larger classroom in order to meet their individualized needs. In addition to differentiated instruction in academic concepts such as reading and writing, some children may need also to be taught the necessary underlying skills for academic success. Such skills may

include planning and goal setting, use of academic language, and basic problem-solving skills.

Experience of Parents in Multicultural Classrooms

Parents who wish to be involved in their child's education may find themselves hampered by several barriers. School climate, usually characterized by how welcoming a school is to its children and their parents, can be a major obstacle to parent-school interactions. If an individual teacher views parental interactions as important, then parents of children in his or her classroom are more likely to become involved. However, a single teacher cannot make an entire school feel welcome; it takes a school-wide effort. Some schools systematically view parental involvement as a burden, whereas other schools integrate parental involvement into all aspects of child learning. If schools and teachers feel that parents are outsiders who are only to be tolerated, then parents sense this and they will stay away. In addition, many culturally diverse families who have had negative experiences in school find it intimidating to be there. In these instances, the school not only needs to be a welcoming place, but it also needs to be a desirable location that eases parental fears of being there.

Poor school climate is only one potential barrier to parental involvement. Another is that many schools are often not effective communicators with multicultural parents. When parents come to school meetings and hear terms such as *individualized education plan, Iowa Tests of Basic Skills*, and *occupational therapist*, the school personnel may as well be speaking only in a foreign language. Many parents just do not understand this type of jargon. This becomes even more complex when the parents really do speak only in a foreign language. If appropriate translators are not available, then these parents are left without any understanding of what is happening with their child's education. Communication difficulties become compounded when correspondence takes place through written communication. The dependence on the use of written communication becomes most problematic when parents cannot read or when they speak a language other than English at home.

In addition to communication problems, parents often do not know what is expected of them in schools. If they come from another country, then they may have no knowledge of what parent-school interactions are supposed to be like in the United States. If they grew up here, but their parents were not involved in their own schools when they were children, then virtually the same experience occurs. Without a clear description of what is expected and desired by the schools, parents can remain stuck in the patterns of their predecessors.

In addition, schools often expect parents to be good decision makers for their children. However, the parents may not share this view. Many parents perceive the teacher as the ultimate power over their child's education. This feeling is especially strong for culturally and linguistically diverse parents. Often, parents believe they will anger teachers if they assert themselves as decision makers for their children. They then worry that angry teachers may victimize their children. These feelings cause them to make few attempts to assert themselves in even the most basic of fashions.

However, rather than be intimidated by their children's school, some parents have chosen instead to take an extremely active role in educational change for their diverse children. Culturally diverse parents led several landmark lawsuits in education. The best known is *Brown v. Board of Education* in 1954, a case brought on behalf of African American children in which the Supreme Court ruled to desegregate schools. It should be noted that this is not the only important case that relates to diversity. *Mills v. the District of Columbia* in 1971 mandated special education for all children, but notably, all of the children represented in that case were African American. *Guadalupe v. Tempe* in 1978 and *Lau v. Nichols* in 1974, cases that established language-specialized educational service for children whose primary language is not English, were brought about by diverse parents on behalf of their children.

Successful Multicultural Teacher-Parent Interactions

Improving school climate helps to overcome many of the cultural barriers between schools and parents. One way for schools to become more inviting includes encouraging families to bring a friend or family member with them to meetings. Another would be to train a school staff member to act as a liaison between the school and the community, thereby allowing families to know at least one person at the school.

Another way to make a school more welcoming is to ensure that the faculty and staff are culturally sensitive. Many teachers are unaware of the multicultural issues affecting parental involvement and thus appear to be insensitive to the needs of the families. School personnel can undertake in service training to learn the skills needed to better serve these children and their families. In addition, teachers can and should learn about different cultures from community members, local churches, and literature searches so they may become aware of potential areas that may make the parent feel disrespected, and smooth the path to better communication. Schools that make parents feel emotionally safe and comfortable encourage parent interactions.

In addition to climate improvements, schools need to be aware of communication problems. Such difficulties occur when parents speak a language other than English, when teachers use jargon that parents cannot understand, or when information is presented in writing to parents who cannot read. Often, it is not the words that are misinterpreted, but the message behind them. Messages should be sent home in a friendly manner that does not make parents feel that the school is speaking down to them because they have less education or less money. All communications should be done with a feeling of respect toward the child, the family, and their culture. Communication issues can be overcome through the use of translators, through simple consideration by teachers, or by presenting information in person or through audio/videotape.

Experience of Teachers in Culturally Competent Classrooms

Of all people involved in the multicultural classroom (teacher, parent, and child), teachers are generally the most comfortable within the school. They understand the system, spend more time there, understand the lingo, and are familiar with the rest of the other faculty and staff. They understand the hidden rules of how to behave in middle class, and they are the ones in charge of the classroom. However, many teachers may not feel as comfortable once they leave the classroom and enter the community surrounding the school.

Given their cultural and economic differences, many teachers who work with ethnically diverse children do not live in or spend time in the neighborhoods of their students. Therefore, they do not share the same daily experiences, know the same environmental landmarks,

or observe the interactions of child and parent on a daily basis. Teachers see the child at school, and they often witness parent-child interactions only in this same location. Based on these observations, they make judgments concerning the identity of their students. However, without a set of shared, daily experiences, they often lack the ability to truly understand all of the strengths, needs, and weaknesses of their students. They need at least some level of understanding of their students' lives to be able to meet their needs.

In addition, although some teachers do an excellent job teaching in the multicultural classroom, others are unprepared to work with diverse students. These teachers may think of a classroom as a melting pot and attempt to teach all children using the same methods. However, the better analogy is that children are more like a tossed salad. Each difference is clearly identifiable, and these differences make the whole more interesting. Thinking that each individual has specific experiences, expectations, and abilities requires the teacher to establish multiple methodologies for working within his or her multicultural classroom. This paradigm shift may be difficult for new teachers to learn or for more experienced teachers to adopt.

Successful Multicultural Classrooms

Cultural and economic learning methodologies among children often translate into differences in learning needs. One technique already discussed is to address students' cultural and economic diversity needs through the use of differentiated instruction. Another is to incorporate the needs of the children and their families into regular classroom policy and philosophy. When teachers and students come from significantly different cultures, sometimes teachers may misinterpret student experiences, abilities, and expectations. Therefore, they may apply teaching and behavior management techniques and strategies that are inappropriate.

In order to learn the appropriate techniques to use, teachers must be sensitive to the needs of their culturally diverse students. To this end, they often attend diversity sensitivity training, where the first step is usually to become aware of their own biases. Following this, teachers are often taught how to become more aware of and appreciative of the differences between themselves and their students. Teachers then use these differences to improve their classroom. Given that teachers' knowledge of cultural diversity varies widely prior to training, diversity instruction

may need to be offered at various levels to meet different teachers' needs.

In addition, teachers need to believe that multicultural differences do not necessarily indicate skill deficiencies for the students. Research on 90/90/90 schools supports the theory that, with some changes within a school, all children can be successful. Reeves defined a 90/90/90 school as one in which at least 90% of the students attending met the criteria for free and reduced lunch, at least 90% of students were non-White, and at least 90% of the students met the required academic standards. Many schools meet these 90/90/90 criteria, and the fact of their existence dispels the notion that ethnically and economically diverse schools cannot also be successful.

Research on these schools indicates that these 90/90/90 schools use a different set of specific strategies than other schools. For example, children in these schools are well aware of what is expected academically and are given instructional interventions to bring them up to these expected levels. They are asked to master reading, writing, and mathematics, and they must apply these skills across subjects. For example, written assignments must be completed in science, math, and social studies classes, not just language arts. These students are assessed frequently, but the assessments are used not to establish accountability, but to determine skill weaknesses. Teachers use these data to work toward helping children master skills instead of demoting or demoralizing the child through constant discussions of poor test performance. Finally, teachers mentor one another in terms of course preparation and grading. They work together to decide what their common goals are and how to meet them.

Implications

In any classroom, there are many differences among children, teachers, and parents. Minor variables such as hair color and eye color have very little impact on learning, but other variables, such as economics and ethnicity, may have considerable impact. Many current instructional strategies often provide more success for White, middle-class children than for economically and culturally diverse ones. However, some schools with large numbers of ethnically and economically diverse children have found success even though the statistics claimed it as improbable. These schools use techniques that address primarily intraschool issues, leading to school performance improvement.

In addition to within-school changes, home-school interactions also need to be altered. Research supports the idea that parent involvement improves academic success. Schools may need to involve their multicultural parents using less traditional methods, such as outreach programs, school liaisons, and just improving school climate. In addition, home-school communication needs to be clear, in the native language, and in multiple formats for those who cannot read in English. Although schools can make positive changes in children's lives without help from parents, more improvements may come about when family involvement is encouraged.

Multicultural schools that have not obtained positive academic outcomes for their students must apply a multifaceted approach to education. Schools should encourage, monitor, and provide opportunities for students; model, consult, and provide training for teachers; and effectively communicate and empower parents to ensure success. Although any one of these techniques should bring about positive results in the multicultural classroom, combinations of all of them will have the strongest impact.

S. Kathleen Krach

See also African Americans; American Indians and Alaska Natives; Asian Americans; Bilingual Education; Bilingualism; Cognitive and Cultural Styles; Cultural Deficit Model; Cultural Diversity; Culture; Diversity; Ethnicity and Race; Hispanic Americans; Home Environment and Academic Intrinsic Motivation; Immigration; Multicultural Education; Risk Factors and Development

Further Readings

Adler, N. (2000). Teaching diverse students. *Multicultural Perspectives*, *2*(2), 28–31.

Delpit, L. (1995). *Other people's children: Cultural conflict in the classroom.* New York: Norton.

Freedman, S., Aschhemi, B., & Zerchykov, R. (1989). *Focus on parents' strategies for increasing the involvement of underrepresented families in education.* Quincy: Massachusetts Department of Education.

Guild, P. (1994). The cultural/learning style connection. *Educational Leadership*, *51*(8), 16–21.

Kalyanpur, M., & Harry, B. (1999). *Culture in special education: Building reciprocal family–professional relationships.* Baltimore, MD: Brookes.

MacMillan, D. L., & Reschly, D. J. (1998). Overrepresentation of minority students: The case for

greater specificity or reconsideration of the variables examined. *Journal of Special Education*, *32*(1), 15–24.

National Center for Education Statistics. (2006). *Digest of education statistics: 2005.* Retrieved October 24, 2006, from http://nces.ed.gov

Payne, R. K. (2005). *A framework for understanding poverty: Fourth revised edition.* Highland, TX: aha! Process.

Reeves, D. B. (2000). *Accountability in action: A blueprint for learning organizations.* Denver, CO: Advanced Learning Press.

Rhodes, R. L., Ochoa, S. H., & Ortiz, S. O. (2005). *Assessing culturally and linguistically diverse students: A practical guide.* New York: Guilford.

U.S. Bureau of the Census. (2006). *Population and household economic topics.* Retrieved October 24, 2006, from http://www.census.gov

Multicultural Education

Multicultural education is a field of study that is defined by various perspectives and disciplines such as education, anthropology, sociology, and psychology. By taking the various definitions and perspectives, multicultural education is best understood as a concept and process designed to help learners recognize, accept, and value differences in culture, ethnicity, social class, sexual orientation, religion, ability, and gender. One of the major objectives of multicultural education is to help all students acquire the knowledge, attitudes, and skills needed to perform effectively in a pluralistic, democratic society. Additionally, students are intended to learn how to interact, negotiate, and communicate with diverse people in an effort to create a civil and moral community. Furthermore, multicultural education aims to promote in learners a sense of responsibility and commitment to work toward democratic ideals of justice, equality, and democracy.

Multicultural education is a concept built on the ideals of freedom, justice, equality, equity, and human dignity as addressed by various documents of the United States such as the Declaration of Independence, the Bill of Rights, and the Constitution.

Multicultural education aims to prepare students for their social responsibilities. Recognizing the influence that schools can have in developing the attitudes and values necessary for a democratic society, multicultural education teaches cultural differences and affirms the pluralism that students, their communities, and teachers reflect. Additionally, it challenges discrimination in schools and society by promoting democratic principles of social justice.

Origins and Foundation

The roots of multicultural education began during the civil rights movements of the United States. Multicultural education is viewed by many as a reaction or response taken by African Americans and other people of color who challenged discriminatory practices in public institutions during the 1960s. Being that educational institutions were among the most oppressive, they were specifically targeted. Activists, community leaders, and parents demanded that educational institutions be more consistent with the racial diversity in the country. Problems such as these were addressed by calling for curricular reform and examining hiring practices in the school system.

During the late 1960s and early 1970s, the women's rights movement began to address issues of education reform as well. Women's rights groups sought to change educational institutions because education was viewed as a contributing factor in institutionalizing sexism. Like groups of ethnic minorities before them, women's rights groups demanded a more inclusive curriculum that covered women's histories and experiences. Additionally, they brought to light the discrepancy between the low numbers of female administrators relative to the percentage of female teachers.

Educational institutions and organizations responded to the concerns brought up by developing a host of programs, practices, and policies. However, these attempts to improve multicultural education mostly led to slight change or addition to the traditional curriculum. The lack of change in educational institutions led those who wanted multicultural education to become dissatisfied with the inequities that were still present in the education system.

At the beginning of the 1980s, education activists and researchers led an emergence of a body of scholarship on multicultural education that no longer tolerated schools addressing their concerns by merely adding token programs and special units on famous women or famous people of color.

By the mid-1980s, other scholars made available more scholarship in multicultural education and developed new and more in-depth frameworks that were based on the ideal of equal educational opportunity and a relationship between school transformation and social change. However, to move beyond slight curricular

changes, the next step was addressing other structural foundations of the school system and the manner in which they contributed to educational inequities. What followed was an exposure, discussion, and criticism of culturally oppressive teaching approaches, standardized tests, and school funding discrepancies; classroom climate; and discriminatory hiring practices.

Goals

Multicultural education sets out to help students develop a positive self-concept through knowledge about histories, cultures, and contributions of diverse groups. Furthermore, it prepares students to work toward attaining and maintaining equality in organizations and institutions by providing knowledge and skills for the redistribution of power and income among diverse groups. Multicultural education advocates that the focus of the teaching and learning process should center on the students' histories and experiences. In line with this method of teaching is the belief that pedagogy should take place in a context that is familiar to students and take into account multiple ways of thinking. Additionally, teachers and students are encouraged to critically analyze oppression and power relations in their communities, society, and the world.

Approach of Reaching Goals

In order to meet the goals set out by multicultural education, it is important that the ideology of multicultural education touch on all aspects of school practices, policies, and organization to ensure the highest levels of academic achievement for all students, regardless of their background or ability. In order to meet the goals of multicultural education, the following is required: a school staff that is culturally competent and racially, culturally, and linguistically diverse to the greatest extent possible. Staff also has to be capable of including and embracing diverse families and communities to create an environment that is supportive of multiple perspectives, experiences, and democracy. Multicultural education requires a complete school reform because it must permeate all aspects of the school system. Taking into account that equality and equity are not the same thing, multicultural education strives to offer all students an equitable educational opportunity while encouraging students to critique society in the interest of social justice.

Miguel Ángel Cano

Further Readings

Banks, J. A., & Banks, C. A. M. (2001). *Multicultural education: Issues and perspectives* (4th ed.). New York: Wiley.

Banks, J. A. E., & Banks, C. A. M. E. (2001). *Handbook of research on multicultural education.* San Francisco: Jossey-Bass.

Hernández, H. (2001). *Multicultural education: A teacher's guide to linking context, process, and content* (2nd ed.). Columbus, OH: Merrill Prentice Hall.

Manning, M. L., & Baruth, L. G. (2000). *Multicultural education of children and adolescents* (3rd ed.). Boston: Allyn & Bacon.

MULTIPLE-CHOICE TESTS

Multiple-choice tests have perhaps the most popular testing format in education and elsewhere, and students are certainly aware of multiple-choice items. Multiple-choice test items begin with the problem typically expressed as an incomplete statement or question (also known as the *stem*) followed by alternatives from which to select (the options).

Certainly, multiple-choice testing is immensely popular in a huge variety of applications in education, psychology, and beyond. Educational psychologists have pioneered in applying multiple-choice testing for research, instruction, assessment/evaluation, prediction/selection/classification/placement, certification/licensure, and for other purposes. Furthermore, educational psychologists have led in conducting research about and development of multiple-choice testing, a short summary of which follows. This entry considers why and how one might use the multiple-choice form, general goals for test developers (including teachers), and more specific item-writing guidelines (especially for multiple-choice items).

Why Use Multiple-Choice Tests?

A multitude of scholars, including Thomas Haladyna and Steven Downing, have indicated a number of strengths of multiple-choice testing. Sampling of content can be excellent, supporting content-valid test interpretations. Reliability of scores can be high. Items

can be easily pretested, scored, used, and reused. Objective, high-speed scoring is possible. Diagnostic subscores are easily obtained. Adaptive testing can be done well with this format. Most content can be tested with this format, including many types of higher-level thinking, even though many writers limit themselves to pedestrian, knowledge-level items. It is a flexible format, and guessing is seldom a problem. If a test-taker is troubled by an item, another item, and opportunity, awaits.

Formal and Informal Uses of Multiple-Choice Tests

Multiple-choice tests are the typical choice to provide big summative evaluations such as whether someone passes a course or earns a professional license. Yet there are also instructional or informal times when a multiple-choice format can encourage student participation, provide corrective feedback to the instructor and students, focus everyone's attention on a topic, give a break from the instructor's droning, preview and prepare for a test, and even allow a microbit of humor.

A Few Goals for Test Developers

Multiple-choice tests should ask important questions consistent with educational objectives or with the projected uses for the test, and ask them clearly. An early National Assessment guideline essentially stated that an item be important and clear enough to be published in the newspaper. Roger Boothroyd concluded that test-takers should either answer correctly for the right reason(s) or, if answering incorrectly, do so for the right reason(s). He is consistent with Lee J. Cronbach's counsel to minimize irrelevant difficulties. Test items are perhaps our most closely read literature and deserve to be based on vital questions that are cleanly stated.

Guidelines for Selecting, Revising, and Writing Good Items

There are three ways to develop test items: selecting existing items, revising existing items, and writing new ones. Educators should evaluate their testing method to be sure that the test items correlate with educational objectives, as some of these methods may prove to be problematic. For example, David Frisbie

and others found that tests accompanying textbooks are likely to include

1. Trivial items, even based merely on phrases lifted from the text,

2. Incomplete coverage,

3. Irrelevant items,

4. Little testing of application, understanding, and critical thinking, and

5. Item faults.

That is, developing a test by appropriating items from a test accompanying the textbook would likely provide a poor sample of items relative to the objectives of the course, with many trivial, irrelevant, low-level items; few items testing higher-level thinking; and big gaps in coverage. Additionally, items may contain errors in item writing, especially item-writing faults that either lead a test-taker to answer correctly for the wrong reason or answer incorrectly for the wrong reason, as cautioned against by Boothroyd and illustrated in Item 5 below.

There are many guides for selecting, revising, and writing quality multiple-choice items, even Robert Ebel's dating back more than 50 years. A more recent taxonomy developed by Haladyna, Downing, and Michael Rodriquez, based on research and textbook authors' advice, is excerpted, modified, and occasionally illustrated in the following paragraphs.

When *selecting item content*, base items on content and mental behavior represented in the table of specifications. Test important and not trivial content. Use novelty to test higher-level learning. Avoid overly specific and overly general content. Avoid trick questions, and keep vocabulary sufficiently simple. A thought-provoking and novel item follows:

1. You are in a boat in a very small pond. You throw four big stones overboard into the water. Assuming precise measurements, the water level in the pond will

A. rise

B. fall

C. stay the same

Solution: The key is B. When the stones are in the boat, the boat displaces water equal to the weight of the stones. When the stones are in the water, the stones displace their volume in water. Which has more volume, 100 lbs. of water or 100 lbs. of stones? So the

stones in the boat displace more water than when they are in the water.

When *formatting*, arrange options vertically. Use multiple-choice, alternate-choice, true-false, multiple true-false, matching, and context-dependent multiple-choice, but avoid complex multiple-choice (Type K), where some choices are combinations of other choices. A Type K item followed by an improvement (three brief, independent items):

2. Which of the following comedians, if any, is/ are noted for verbal attacks on others?

    A. Bill Cosby

    B. Don Rickles

    C. Red Skelton

    D. A & B

    E. A & C

    F. B & C

    G. A, B, & C

    H. None of the above.

2–4. For each of the following comedians, indicate whether he is noted for verbal attacks on others. Use *A* if *yes*, *B* if *no.*

    2. Bill Cosby

    3. Don Rickles

    4. Red Skelton

Concerning *style*, edit and proof. Use correct grammar. Minimize the amount of reading test-takers must do.

When writing the stem (the incomplete statement or question specifying the problem to be solved), include the item's central idea. Avoid extra verbiage. Move repeated words to the stem. Word the stem positively, but if negative phrasing is necessary, capitalize or bold the negative word. To illustrate,

5. Validity is not

    A. related to reliability.

    B. an appropriate concern for test users.

    C. synonymous with reliability, stability, equivalence, or internal consistency.

    D. involved when domain referencing.

    E. important.

This item, somewhat similar to one found in a set of items accompanying a measurement text, starts off poorly. The stem does not specify the problem addressed by the item, which would have helped direct the item writer and the test-taker; it contains a negative, and the negative is not bolded or in caps.

When writing *choices*, develop as many effective choices as you can, but three is likely fine. Ensure that only one choice is correct or best. Order choices logically, numerically, or alphabetically. Keep choices homogeneous in content and grammar. Make choice lengths about equal. Use "none of the above" carefully. Eschew "all of the above." Avoid clues such as specific determiners (e.g., *always, absolutely*), choices clued by the stem, grammatical inconsistencies, and pairs or triplets of options that clue the correct choice. (Inattention to some of these guidelines is illustrated in Item 5 above: At least one choice is implausible, so the number of options could be trimmed; choices are not alphabetized, so the key is C; choices are not homogeneous in content or grammar; the best answer is the longest and contains the word *synonymous;* and two options contain the same concept, so the key is likely A or C.) Two other guidelines: Use typical errors to write distractors (i.e., the wrong alternatives; the alternatives that are *not* the correct or best alternative). Incorporate humor if compatible with the teacher and situation. Here is an item with a bit of humor:

6. Directions: Select the letter corresponding to the line containing any grammatical error.

    A. the umpire's new glasses seemed to

    B. be helping him until he called

    C. a eagle that flew by a foul ball

    D. (no mistakes)

Multiple-choice tests, carefully crafted, can have many strengths. Furthermore, as Haladyna discussed, attempts are being made to automate item writing, to apply cognitive science, to integrate validity theory, and to analyze distractor performance. Perhaps the challenge for all of us is best expressed by various cartoonists' captions, such as, "I don't yet have all the answers, but I am beginning to ask the right questions."

Robert F. McMorris and Xiaoyuan Tan

See also Aptitude Tests; Bloom's Taxonomy of Educational Objectives; Certification; Criterion-Referenced Testing;

High-Stakes Testing; Measurement; Norm-Referenced Tests; Standardized Tests; Testing

Further Readings

Berk, R. A. (2002). *Humor as an instructional defibrillator: Evidence-based techniques in teaching and assessment.* Sterling, VA: Stylus.

Carter, K. (1986). Test-wiseness for teachers and students. *Educational Measurement: Issues and Practice*, *5*(4), 20–23.

Ebel, R. L. (1951). Writing the test item. In E. F. Lindquist (Ed.), *Educational measurement* (pp. 185–249). Washington, DC: American Council on Education.

Frisbie, D. A., Miranda, D. U., & Baker, K. K. (1993). An evaluation of elementary textbook tests as classroom assessment tools. *Applied Measurement in Education*, *6*(1), 21–36.

Haladyna, T. M. (2004). *Developing and validating multiple-choice test items.* Mahwah, NJ: Lawrence Erlbaum.

Haladyna, T. M., & Downing, S. M. (1989). A taxonomy of multiple-choice item-writing rules. *Applied Measurement in Education*, *2*, 37–50.

Haladyna, T. M., Downing, S. M., & Rodriguez, M. C. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. *Applied Measurement in Education*, *15*(3), 309–334.

McMorris, R. F., & Boothroyd, R. A. (1993). Tests that teachers build: An analysis of classroom tests in science and mathematics. *Applied Measurement in Education*, *6*, 321–342.

Multiple Intelligences

When Howard Gardner introduced the theory of multiple intelligences (MI) in the early 1980s, he proposed a departure from the conventional view of intelligence. In MI theory, Gardner (1999) defined intelligences as "a biopsychological potential to process information that can be activated in a cultural setting to solve problems or create products that are of value in a culture" (pp. 33–34). In contrast to *g*, a measure of general intelligence that is typically assessed using standardized intelligence tests or achievement tests (such as the SAT or ACT), MI theory has delineated at least eight intelligences that all people have available to them, and use, throughout their lives. It is important to note that, although everyone has access to all of the multiple intelligences, most people have strengths in one or two particular abilities.

The eight intelligences are as follows:

1. Linguistic intelligence,
2. Mathematical/logical intelligence,
3. Bodily/kinesthetic intelligence,
4. Interpersonal intelligence,
5. Intrapersonal intelligence,
6. Naturalistic intelligence,
7. Musical intelligence, and
8. Spatial intelligence.

Linguistic intelligence and mathematical/logical intelligence are the two intelligences assessed using standardized assessments such as the SAT or ACT, intelligence tests such as the Wechsler Adult Intelligence Scale (WAIS) and the Wechsler Intelligence Scale for Children (WISC), and the Wechsler Preschool and Primary Scale of Intelligence (WPPSI). Gardner suggested that the emphasis on linguistic and mathematical/logical abilities is one of the biases of conventional ideas about intelligence. In contrast, MI theory proposes that some abilities, such as interpersonal intelligence or naturalistic intelligence, are not as readily measurable but are still important abilities that ought to be acknowledged and nurtured to strengthen human development.

In MI theory, eight criteria must be met for a strength or ability to be considered one of the multiple intelligences: (1) in the brain, the location of an intelligence can be isolated by brain damage; (2) its evolutionary history can be traced and is plausible; (3) its core operations must be readily identifiable; (4) there is a symbol system associated with the ability (such as the alphabet in linguistic intelligence); (5) its developmental history can be traced and a set of optimal performances has been identified; (6) idiot savants, prodigies, and other extraordinary people have been identified; and both (7) experimental psychological tasks and (8) psychometrics evidence support its existence. The following sections summarize each of the eight intelligences meeting the inclusion criteria set forth by MI theory. In addition, examples of careers that are well-suited for individuals who have strengths in a given intelligence are discussed. Activities that can be used to strengthen each of the intelligences also are provided.

Linguistic Intelligence

Linguistic intelligence is characterized by a sensitivity to languages, including the ability to learn languages and to use languages to achieve goals. It is one of the

later-developing intelligences, because to be linguistically adept requires life experience. Early talkers and readers may not become linguistically gifted if their life experiences are limited. Four components comprise linguistic intelligence: phonology, syntax, semantics, and pragmatics. Linguistically intelligent people are aware of the phonetics, or sounds, of language. As a result, they usually possess an advanced verbal sense of humor—they often use language to make puns, analogies, tongue twisters, and jokes. Those with strong linguistic abilities are able to manipulate the syntax, or structure, of language. They are able to compose exceptionally structured paragraphs and may be particularly sensitive to the composition and grammar in their own and others' oral and written language. Foreign languages come easily, as do reading and memorization. The linguistically gifted tend to be sensitive to semantics, paying close attention to the subtleties of words and their meaning. They are pragmatic in their use of language; they entertain, persuade, teach, and lead through the written and spoken word. Broca's area is the region of the brain that is associated with linguistic intelligence. When Broca's area is damaged, patients are able to comprehend language but are typically unable to generate even simple sentences. In most cultures, language is subject to symbolic encoding (i.e., the alphabet and the written word), thus meeting one of MI's criteria for inclusion.

Examples of linguistically gifted people include Virginia Woolf, whose literary work *A Room of One's Own* spoke to her concerns for women's rights to financial freedom and privacy; and Mark Twain, whose humor and satire have come to characterize American literature. The career paths for linguistically gifted people include creative and scientific writing, performance in the media and onstage, politics and law, teaching at any educational level, and language translation. Activities that can be used to train linguistic intelligence include listening to recordings of famous speakers, reading one book per week, subscribing to a literary magazine, and memorizing poetry and prose.

Mathematical/Logical Intelligence

Mathematical/logical intelligence is distinguished by the ability to solve problems logically, complete mathematical problems quickly, differentiate logical or numerical patterns, and conduct scientific inquiry. To those gifted in mathematical/logical intelligence, the problem-solving process is typically rapid and nonverbal. Solutions often present themselves "out of the blue" and may seem to be invisible to the problem solver. Along with linguistic intelligence, mathematical/logical abilities are at the core of intelligence testing. Areas of the brain involved with mathematical/logical intelligence include the frontotemporal lobes (logic) and the parietofrontal lobes (numerical calculation). From a developmental perspective, child prodigies and savants exist, providing additional support for the inclusion of mathematical/logical intelligence in MI theory. Furthermore, Jean Piaget's work documented the developmental path of mathematical/logical abilities. Mathematical operations provide a key example of an intelligence that is subject to an encoding system.

Mathematician John Nash, who was awarded the Nobel Prize in Economics for his analysis of game theory, and astronaut Katherine Sullivan, the first woman to walk in space, are examples of people who are extraordinarily gifted in mathematical/logical intelligence. Career paths for individuals gifted in mathematical/logical intelligence include mathematics and statistics, philosophy, physics, chemistry, biochemistry, engineering, and computer programming. For those strong in mathematical/logical intelligence and linguistic intelligence, careers in scientific communications, research and teaching at the university level, and pharmaceutical drug development may be appropriate. Activities that can be used to train mathematical/logical intelligence include playing math and logic games and brain teasers, learning a computer language, and teaching math and science concepts.

Bodily/Kinesthetic Intelligence

Bodily/kinesthetic intelligence is distinguished by the ability to expertly control one's body movements and the ability to skillfully handle objects. People who have excellent bodily/kinesthetic intelligence are able to use their hands or their bodies to create items and to solve problems. From an evolutionary perspective, the specialization of body movements has allowed humans to adapt; the development and use of tools further advances the evolution of the human species. In the brain, the motor cortex is responsible for the control of bodily movement. In right-handed people, motor control is located in the left hemisphere; in left-handed people, motor control is located in the right hemisphere. Brain damage, such as apraxia (loss of the ability to carry out movements despite having

the willingness and the ability to perform the movements), provides evidence for the existence of bodily/kinesthetic intelligence. Piaget characterized the qualities associated with this intelligence as the sensorimotor stage of cognitive development. In fact, children's body movements, such as balance, coordination, dexterity, flexibility, reflexes, strength, and expressiveness, develop according to a well-characterized schedule. Those with exceptional bodily-kinesthetic intelligence possess the ability to imagine the movement of their bodies in space and to carry out the actions in physical form.

End-state expertise of bodily-kinesthetic intelligence is illustrated by dancer and choreographer Martha Graham, a pioneer of modern dance, and by Brian Boitano, American figure skating champion. Careers suited for those with bodily-kinesthetic intelligence include actors, stunt people, comedians, dancers, horse trainers, athletes, coaches, figure skaters, and artisans. The following activities can be performed to train bodily-kinesthetic intelligence: taking lessons in a sport such as golf, swimming, gymnastics, or dance; playing charades; taking acting classes; practicing yoga; playing video games that require quick reflexes; and engaging in imagery and visualization experiences that emphasize bodily movements.

Interpersonal Intelligence

Interpersonal intelligence is characterized by the ability to accurately evaluate the moods, intentions, thoughts, feelings, and motivations of other people. From an evolutionary perspective, interpersonal intelligence affects the ways that people interact with and sustain communities. The interpersonal qualities of leadership, teaching, and healing all take place within a social context and play a role in the survival and evolution of societies. The frontal lobe is indicated as the area of the brain involved with interpersonal intelligence. Damage to this area via injury, lobotomy, or Pick's disease (a type of dementia) results in lasting personality changes, although other problem-solving capacities are left intact. Evidence for a biological basis for interpersonal intelligence has been linked to two factors. During the prolonged childhood of primates, attachment to a mother seems to be an important aspect of normal interpersonal development. In cases when the mother figure is not present or involved with her young, normal development of primates has been shown to be compromised. Next, in

humans, social interactions such as cooperation, leadership, group cohesion, and organization have played salient roles in completing skilled tasks such as hunting, tracking, and the formation and maintenance of communities. Facial expressions, gestures, other body language, and verbal cues make up the symbol system for interpersonal intelligence.

Eminent leaders, such as civil rights activist Martin Luther King, Jr., and humanitarian Oprah Winfrey, are among those individuals who demonstrate the existence of exceptional interpersonal intelligence. Careers suited for those with interpersonal intelligence include teaching, spiritual and political leadership, sales, and acting. The following activities can be used to develop interpersonal intelligence: practice evaluating other people's moods and feelings, teach or counsel people through a volunteer service organization, go to a public place and observe how people interact, meet a new person every day, take a course that focuses on interpersonal communications, make eye contact with others during conversations, and notice people's nonverbal communication (body language) as they interact.

Intrapersonal Intelligence

Intrapersonal intelligence is characterized by the ability to access and understand one's own internal experiences, including a range of emotions, and to draw on internal experiences as a means of making decisions about, and guiding, one's behavior. People with strong intrapersonal intelligence also have an accurate understanding of how they fit in relation to other people and have a strong sense of self. They have the ability to be both creative and intimate, and they possess the capacity to be alone. Because this particular intelligence is psychically manifested, language, music, or other creative expression is often used to describe the experiences of an intrapersonally intelligent person.

As is the case with interpersonal intelligence, the frontal lobe is involved with personality. Damage to the frontal lobe can alter personality but leave other cognitive functions intact. An individual may be irritable or euphoric in the case of damage to the lower areas of the frontal lobe; or listless, slow, and apathetic if damage occurs in the higher regions. The development of intrapersonal intelligence seems to start in early childhood, as children learn to understand their identities in the context of the world around them. A positive self-concept tends to develop when children

are nurtured and loved. Autistic individuals provide an example of those in whom intrapersonal intelligence is impaired. Although an autistic individual may not be able to refer to him- or herself, he or she typically demonstrates extraordinary abilities in mathematical/logical intelligence or in musical intelligence. Gardner acknowledges that evolutionary evidence of intrapersonal intelligence is difficult to identify. However, intrapersonal intelligence is thought to serve the purpose of assisting people to overcome or manage their basic human instincts after the need for survival has been met. As the symbol system that most closely maps to intrapersonal intelligence, dreams offer symbols that relate to aspects of the self. Finally, intrapersonal intelligence seems to play a role in the advancement of culture, as this particular ability helps people better understand themselves and perhaps proceed to make choices that benefit society.

The fourteenth Dalai Lama, Tenzin Gyatso, is the spiritual and political leader of the Tibetan people and an example of an individual with exceptional intrapersonal intelligence. Through his speeches, writings, and actions, he demonstrates an attitude of peace, harmony, and nonviolence for which he was awarded the Nobel Prize for Peace in 1989. Similarly, Mother Teresa, leader of the Missionaries of Charity Calcutta, felt she was called by God to become a missionary and eventually started an open-air school for the poorest of the poor in the slums of Calcutta. Her self-awareness and devotion to her calling provide another example of an individual with excellent intrapersonal intelligence. Because intrapersonal intelligence is expressed through other intelligences, the career paths for people with well-developed intrapersonal intelligence are varied and include spiritual and political leadership, psychology, psychotherapy, and teaching. To train intrapersonal intelligence, people can engage in counseling or psychotherapy, learn and practice meditation, keep a dream journal, and read self-help books and other inspirational books.

Naturalistic Intelligence

Naturalistic intelligence is characterized by the ability to recognize and classify different species of flora and fauna in the environment. From an evolutionary perspective, naturalistic intelligence ensured the survival of the human species. Hunters and gatherers necessarily had to distinguish sustainable nutrients from poisonous varieties, and differentiate predators from prey.

Naturalistic intelligence has been subject to encoding and language throughout history: taxonomic and linguistic systems exist for encoding varieties of species. It is unclear which neural centers are involved in naturalistic intelligence. However, in brain-damaged patients, the ability to recognize and classify living organisms is absent, whereas the ability to name and recognize inanimate objects remains intact. The opposite case has been found as well. Some brain-damaged patients retain the ability to name and recognize living objects, whereas they cannot do the same with inanimate objects. This phenomenon provides initial support for the likelihood that abilities associated with naturalistic intelligence may be localized in the brain. Developmental evidence for naturalistic intelligence consists of the observation that children are frequently interested and involved with the identification and naming of objects in the natural world. For example, children may display an enthusiasm for dinosaurs or insects; they may be found digging in the dirt, collecting rocks, and playing in the trees. Later on, naturalistic intelligence may manifest as specific abilities in gardening, hunting and fishing, and hiking and camping, as well as involvement in environmental organizations.

Regarding end-state expertise, naturalist Charles Darwin is an example of an individual who is gifted in naturalistic intelligence. His work to identify and classify species of fossils, plants, and animals became the basis of his theory of natural selection and culminated in his seminal work *On the Origin of Species.* Primatologist Jane Goodall is another example of naturalistic intelligence in action. Goodall's research illustrates her efforts to protect and understand the complex relationships of chimpanzees. Professionally, people with naturalistic intelligence may pursue careers in gardening, environmental science, botany, entomology, veterinary medicine, viticulture (wine making), and wildlife management and preservation. Activities that can be employed to train naturalistic intelligence include spending time in the natural world, studying species of birds and other wildlife, reading magazines that focus on nature, and keeping a nature journal that details observations and questions about the natural world.

Musical Intelligence

Musical intelligence is characterized by auditory imagery (the ability to mentally hear musical tones), and the ability to hear, recognize, and manipulate

music. From an evolutionary perspective, the use of music to unify people can be traced to Stone Age societies. The right hemisphere of the brain seems to be associated with musical perceptions and production. However, musical abilities are not clearly localized to a specific region of the brain. The presence of savants suggests that there is a biological predisposition for extreme musical abilities. Musical notation gives evidence that music is subject to an encoding system. From a developmental perspective, evidence suggests that infants can demonstrate computational abilities. Even before they are able to play musical instruments, some young children are adept with rhythm and sound, an occurrence that provides developmental evidence for musical intelligence. Later, children with musical abilities quickly master musical instruments and can demonstrate perfect pitch.

End-state expertise in musical intelligence is evidenced in child prodigies who move quickly to an advanced level of performance. For example, Mozart's musical abilities became apparent around the age of 3. By age 4, he had mastered several pieces for the piano and composed his first pieces at age 5. Midori is a contemporary example of extreme musical ability. Attracted to violin music at age 3, she was given her first violin 1 year later. Midori's first public performance as a violinist took place when she was 7 years old. Professionally, people with musical abilities may choose careers as performers, instrument tuners and instrument builders, and music educators. Musical intelligence can be developed by attending concerts and musical performances, taking music lessons, participating in a choir or band, and listening to unfamiliar genres of music.

Spatial Intelligence

Spatial intelligence is characterized by the ability to recognize and manipulate areas of space. Specifically, people who possess spatial intelligence have the capacity to perceive the world accurately and are able to transform their initial perceptions through mental rotation. They often see things that other people miss and apply their spatial abilities to arts such as sculpture, invention, painting, photography, interior design, and architecture. Other applications of spatial intelligence include navigation, reconnaissance, and piloting aircraft. The posterior region of the right cerebral cortex is the location in the brain where spatial processing occurs. People who sustain brain damage in this region have difficulty finding their way around a location, recognizing faces, and noticing details. Spatial intelligence contributes to the evolution of a variety of domains, from navigation to sculpture and other visual arts.

Examples of people who possess an end-state expertise in spatial intelligence include Frank Lloyd Wright, whose organic style of architecture emphasized balance between the natural world and the needs of human inhabitants, and Frieda Kahlo, noted Mexican painter, whose artistic style combines surrealism, realism, and symbolism. Careers for those with spatial intelligence include pilot, navigator, sculptor, mechanical engineer, architect, photographer, computer graphics designer, interior designer, and other careers in the visual arts. Spatial intelligence can be developed by studying geometry, learning photography, studying optical illusions, and learning navigation skills.

Evidence for Spiritual and Existential Moral Intelligences

In the decades following the introduction of MI theory, candidate intelligences have been considered for inclusion in MI theory. The other abilities, including spiritual and existential, have been subjected to the criteria set forth in MI theory. Gardner's assessment is that none of these human abilities meets all eight of the requirements to be included as a multiple intelligence. Gardner does, however, suggest that an aspect of spirituality, namely existential intelligence, does hold up to the eight criteria fairly well. Existential intelligence is concerned with the "big questions" in life and deals with the existence of God and the meaning of life. Yet the lack of biological evidence prevents existential intelligence from being included.

Robyn McKay

See also Creativity; Individual Differences; Intelligence and Intellectual Development; Intelligence Tests

Further Readings

Armstrong, T. (1999). *7 kinds of smart: Identifying and developing your multiple intelligences.* New York: Plume.

Armstrong, T. (2000). *Multiple intelligences in the classroom.* Alexandria, VA: Association for Supervision & Curriculum Development.

Baum, S., Veins, J., & Slatin, B. (2005). *Multiple intelligences in the elementary classroom: A teacher's toolkit.* New York: College Teachers Press.

Beachner, L., & Pickett, A. (2001). *Multiple intelligences and positive life habits: 174 activities for applying them in your classroom.* Thousand Oaks, CA: Corwin.

Gardner, H. (1999). *Intelligence reframed: Multiple intelligences for the 21st century.* New York: Basic Books.

Gardner, H. (2006). *Multiple intelligences: New horizons.* New York: Basic Books.

Nelson, K., & Nicholson-Nelson, K. (1999). *Developing students' multiple intelligences (grades K–8).* New York: Scholastic.

MYELINATION

Myelination, also known as myelinization, is a process that occurs in the brain wherein a myelin sheath is formed around a nerve fiber. Myelinated neurons are the white matter of the brain. The explosion of neurological research has increased educational psychologists' interest in exploring how the brain works and its relationship to learning as myelination is a developmental process. Important issues of myelination include how and when it is formed and what new research says about its production and destruction.

Myelin is composed of lipid fat and protein. It is produced by different cell types and varies in its chemical composition. However, although varied by cell of origin and composition, all myelin performs the same function—insulation. The process of myelination occurs when the myelin envelops the axons in the brain, thereby connecting neuron to neuron and making the brain into a singular whole. The main function of myelination is insulation and therefore it dramatically increases the speed by which electrical impulses are transmitted between nerve cells in the brain. This insulation also helps to prevent signals from leaving the axons. An additional function of myelination is the protection of axons from external danger. When a peripheral fiber is severed, the myelin sheath can help to provide a track along which regrowth can occur between axons.

Myelination is a developmental process that occurs in the brain. In fact, recent research suggests that brain development occurs throughout the lifespan of humans. The brain is composed of many regions with specific functions. The maturation of these regions and their connecting pathways is required for one to be successful in the development of cognitive, motor, and sensory functions. As the brain matures and connections are made between regions, the smooth flow of information is increased by the process of myelination. This extensive process of myelination increases information processing speed and underlies the processes of language and executive functions of the brain.

Myelination begins during the fifth fetal month with the myelination of the cranial nerves. Postmortem studies suggest that the myelination process is extensive during the first 2 years of life. The first 10 years of life are dedicated to the myelination in regions of the brain primarily responsible for vision, hearing, and motor skills as well as the areas responsible for basic thinking skills. However, the previous notion that brain development ends in childhood is disproved by recent research that instead indicates that the brain may not be fully mature into late adulthood, maybe even the entire life span. Although the timetable of myelination during childhood and adolescence is still being researched, the current research suggests a gradual maturation during late childhood and early adolescence of fiber pathways that support motor and speech functions. Portions of the brain responsible for higher thinking skills and abstract or reasoned thought appear to be myelinated in the 20s and 30s or even later in life. Maximum myelination is thought to occur around the age of 50.

The myelination process is vulnerable to genetic and environmental factors, including toxins, genetic predispositions, and even diet. The disruption of myelination is a key neurobiological component that makes the human brain susceptible to developmental disorders such as autism, learning disabilities, attention deficit hyperactivity disorder, schizophrenia, and addiction. Also, decay of myelin due to autoimmune damage is suspected in diseases such as multiple sclerosis. The midlife breakdown of myelin could be a key to the onset of Alzheimer's disease as well.

More research is needed to determine what affects the process of myelination. Research continues to look at the role of diet and nutrition in the process of myelination. Also, the development of drugs that mimic the stimulation of cells that produce myelin may be effective in the treatment of demyelinating diseases such as multiple sclerosis. It is crucial that more is known about how myelination takes place to offer insight into the learning process as myelination determines the character and nature of one's mind and personality.

Lori Jackson

See also Attention Deficit Hyperactivity Disorder; Learning Disabilities

Further Readings

Bartzokis, G. (2006). Brain myelination in prevalent neuropsychiatric developmental disorders: Primary and comorbid addiction. *Adolescent Psychiatry*, *29*, 55–96.

Ishibashi, T., Dakin, K. A., Stevens, B., Lee, P. R., Kozlov, S. V., Stewart, C. L., et al. (2006). Astrocytes promote myelination in response to electrical impulses. *Neuron*, *49*, 823–832.

Paus, T., Zijdenbos, A., Worsley, K., Collins, D. L., Blumenthal, J., Giedd, J. N., et al. (1999). Structural maturation of neural pathways in children and adolescents: In vivo study. *Science*, *283*, 1908–1912.

N

A three-year-old child is a being who gets almost as much fun out of a fifty-six-dollar set of swings as it does out of finding a small green worm.

—Bill Vaughan

NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS

The National Assessment of Educational Progress (NAEP), also known as "The Nation's Report Card™," is a nationally representative and ongoing evaluation of how America's students perform across various subject areas. The NAEP focuses on two major goals: to compare student achievement in states and to chart changes in achievement of 4th, 8th, and 12th graders across time in mathematics, reading, writing, science, and other content domains. The NAEP, even today, constitutes the only continually conducted assessment that represents the academic status of all students in the nation.

Congressionally mandated by the National Center for Education Statistics (NCES), the NAEP project is housed within the Institute of Education Sciences (IES) of the U.S. Department of Education. Oversight is provided by the National Assessment Governing Board (NAGB). The NAGB is responsible for setting policy for NAEP as well as for specifying the framework under which assessments will be developed and conducted. NAGB, a 26-member board, includes governors, state legislators, local and state school officials, educators, business representatives, and members of the general public.

The NAEP incorporates several components: National NAEP, State NAEP, NAEP Trial Urban District Assessment, Long Term Trend, and Special Studies. The National NAEP reports information by state and nationally, and provides results for students in 4th, 8th, and 12th grade. The State NAEP provides state-specific results designed to support accurate and representative state-level results. The NAEP Trial Urban District Assessment provides district-specific results for large urban school districts (e.g., Atlanta City School District, Boston Public Schools). The Special Studies target contemporary measurement/assessment questions.

One of the primary functions of the NAEP is to chart student performance trends over time. Long-term assessments make it possible, decade after decade, to measure student progress in reading and math. NAEP findings also provide an objective annual picture available to various stakeholders (e.g., national, state, and local policymakers) on the condition and progress of education in the United States. Subject matter achievement is reported in two ways: scale scores and achievement level. NAEP scale score results provide a numeric summary of what students know and can accomplish in a particular subject and are presented for groups (e.g., gender) and subgroups (e.g., ethnicity). Achievement levels categorize student achievement as Basic, Proficient, and Advanced, using ranges of performance established for each grade.

NAEP results are drawn from random samples of student populations. Results of these assessments are not provided to individual students or schools. Instead, results are incorporated into a general report involving subject matter achievement, instructional experiences, and school environment of varied populations of interest (e.g., third graders) as well as groups within these populations (e.g., male students, African American students). The main NAEP assessments are conducted at the national and state levels every 2 years.

Since 1990, NAEP assessments have also been conducted at the state level for states wishing to participate. Although the content of state and national assessments is identical, student samples were formerly drawn separately for both levels. Since 2002, a combined sample has been selected. Typically, in each state, approximately 2,500 students across 100 public schools are randomly selected per grade, per subject. State-level assessments are conducted at Grades 4 and 8, but not 12. Like the national assessment, State NAEP does not provide individual student or school scores. Participation allows states the added benefit of monitoring their own progress over time in the selected subject areas.

To ensure that the national sample is representative of the total national student population, states that decide not to participate at the state level are still asked to participate at the national level. Results for schools in those states are identified only as "national." This makes the national sample a combination of all samples from participating states plus all samples from nonparticipatory states.

The NAEP long-term trend assessments are administered across the nation every 4 years to determine patterns of change in the academic performance of American students. Results of these evaluations are reported for students ages 9, 13, and 17 in mathematics and reading. Although they both evaluate mathematics and reading, long-term trend assessments are different from the national and state assessments. For instance, unlike national and state assessments, which evolve with each decade to keep up with current school practices, long-term trend assessments remain constant. This is done in an effort to replicate procedures as precisely as possible in order to measure change effectively over time.

Since 2002, the NAEP project has been conducting trial district-level assessments in certain urban districts. In 2002, participating districts were Atlanta, Chicago, Houston, Los Angeles, New York City, and the District of Columbia. In 2003, nine districts—Atlanta City School District, Boston School District, Charlotte-Mecklenburg Schools, City of Chicago School District 299, Cleveland Municipal School District, Houston Independent School District, Los Angeles Unified School District, New York City Public Schools, and San Diego Unified School District—were selected for assessment. In 2005, the same school districts tested in 2003, plus Austin Independent School District, were evaluated. Results for these districts are also compared with results for public school students in large central cities and the nation.

NAEP also coordinates various studies related to assessment. One ongoing project, the Technology-Based Assessment, investigates the use of technology, particularly computers, as tools that aid the quality and efficiency of educational assessments. Other projects include the oral reading study (2002) and the National Indian Education Study (2005).

Jorge E. Gonzalez and Sophia Tani-Prado

See also Assessment; Evaluation

Further Readings

Perie, M., & Moran, R. (2005). *NAEP 2004 trends in academic progress: Three decades of student performance in reading and mathematics* (NCES 2005–464). Washington, DC: U.S. Government Printing Office.

Web Sites

The IES National Center for Education Statistics: The Nation's Report Card: http://nces.ed.gov/nationsreportcard

NATIONAL CENTER FOR EDUCATION STATISTICS

The National Center for Education Statistics (NCES), part of the U.S. Department of Education and the Institute of Education Sciences, is the primary federal entity for generating and analyzing national and international data used to examine the American educational system. More specifically, the mission of the NCES is to collect, collate, analyze, and report statistics on the condition of education in the United States, as well as to publish reports and analyses explaining the results of statistical findings. The NCES uses the

latest advances in survey design and statistical methods, presents findings in a variety of formats, and provides data that address educational issues of current interest. Through the collection of data and the presentation of results in a variety of publications, including annual reports, information on the status of the educational system in the United States is available for use by researchers, professional educational organizations, legislators, and the general public.

Data Collection

The NCES collects data primarily through the administration of surveys covering a variety of topics, including educational assessments, early childhood experiences, elementary and secondary education, international educational issues and trends, and postsecondary education.

Educational Assessments

The NCES collects and reports information on the academic performance of students in the nation's schools and the literacy skills of the nation's adults. The National Assessment of Educational Progress (NAEP), also known as the "Nation's Report Card," is the only nationally representative and continuing assessment of the academic achievement of students in various subjects. Through the National Assessments of Adult Literacy (NAAL), basic literacy and mathematical skills of adults are assessed, and with this information, the status of adult literacy, national trends in adult literacy, and relationships between literacy and adult characteristics are revealed.

Early Childhood

The NCES collects data on children's health and early care school experiences through two longitudinal studies: the Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K) and the Early Childhood Longitudinal Study, Birth Cohort (ECLS-B). These studies are two of the first nationally representative studies of children's developmental and educational experiences from birth to early adolescence.

Elementary/Secondary

The NCES collects data on the condition of public and private education through the Common Core of Data (CCD), a database that holds information on the nation's public elementary and secondary schools (approximately 94,000) and school districts (approximately 17,000); the Schools and Staffing Survey (SASS), which collects information on principals, teachers, schools, school districts, and library media centers; the Private School Survey (PSS), which compiles a list of the nation's private schools; and the National Household Education Survey (NHES), which is used to collect information on a variety of educational issues—for example, in 2003, surveys focusing on Adult Education for Work Related Reasons and Parent and Family Involvement in Education were administered.

International

The International Activities Program provides data comparing the educational experiences of students and educational trends in the United States to those in other countries. Trends in mathematics and science achievement, literacy achievement, and civics are among the issues examined by these surveys.

Postsecondary

Through the administration of the Integrated Postsecondary Education Data System (IPEDS) Survey, a variety of data on the nation's 9,800 public and private postsecondary institutions, including enrollments, program completions, faculty, staff, finances, and academic libraries, are collected. Additional surveys collect data on students, financial aid, postsecondary faculty, degree recipients, and transcripts.

Annual Reports

In addition to the statistical information collected, the NCES presents findings in a variety of formats, including five main annual reports: *The Condition of Education, The Digest of Education Statistics, Indicators of School Crime and Safety, Projections of Education Statistics*, and *Education Statistics Quarterly.*

First, *The Condition of Education* summarizes development and trends in education using the most recent available data. This report includes findings from six main areas: (1) enrollment trends and student characteristics at all levels; (2) student achievement and the long-term effects of education; (3) student effort and rates of progress through the educational system; (4) the contexts of elementary and secondary

education; (5) the contexts of postsecondary education; and (6) support from society for learning, including parental and community support, and public and private financial support of all levels of education.

Second, *The Digest of Education Statistics* compiles statistical information on topics of current interest covering a wide range of educational issues at all levels. Issues discussed include the number of schools, colleges, teachers, enrollments, and graduates as well as educational attainment, federal funds for education, libraries, employment and income of graduates, and international educational comparisons.

Third, *Indicators of School Crime and Safety* is published jointly by the NCES and the Bureau of Justice Statistics. It examines crime occurring on the way to and from school, as well as in school. Through the National Crime Victimization Survey, the School Crime Supplement to the National Crime Victimization Survey, the Youth Risk Behavior Survey, the School Survey on Crime and Safety, and the Schools and Staffing Survey, data on the nature of crime in schools are collected and presented from the perspectives of students, teachers, principals, and the general population.

Fourth, *Projections of Education Statistics* provides projections on major education statistics, such as statistics on elementary and secondary schools' enrollment, graduates, teachers, and expenditures, as well as enrollment, degrees earned, and expenditures of institutions of higher education. Data from the past 15 years, as well as projections for the future, are displayed in text, tables, and figures in the publication.

Finally, *Education Statistics Quarterly* is a comprehensive overview of all NCES projects. Each issue includes short publications and summaries of all NCES publications and data products recently released. Furthermore, training and funding opportunities are included, as well as a featured topic with commentary.

Carolyn L. Carlson

See also Descriptive Statistics; Longitudinal Research; National Assessment of Educational Progress

Further Readings

Forgione, P. D. (1999). *Using federal statistics to inform educational policy and practice: The role of the National Center for Education Statistics* (presented to the U.S. Senate Committee on Health, Education, Labor, and Pensions). Washington, DC: Office of Educational Research and Improvement.

Snyder, T. D., Tan, A. G., & Hoffman, C. M. (2006). *Digest of Education Statistics*, *2005* (NCES 2006–030). Washington, DC: U.S. Government Printing Office.

U.S. Department of Education, National Center for Education Statistics. (2006). *The Condition of Education 2006* (NCES 2006–071). Washington, DC: U.S. Government Printing Office.

Web Sites

National Center for Education Statistics: http://www.nces.ed.gov

NATURALISTIC OBSERVATION

Naturalistic observation refers to the scientific study of the social interactions of children and adults within their native environments, such as in school or at home, or within artificial situations designed to elicit specific behaviors of interest that occur in native environments, such as child social and problem-solving skills with peers or parents. In educational psychology, the phrase is most commonly used to describe studies in which researchers or practitioners attempt to observe subjects of interest as unobtrusively and impartially as possible. Observation techniques were developed throughout the 20th century but came into prominence with the rise of behavior therapy and behavioral assessment in the 1960s. Naturalistic observation is particularly useful for describing the social, physical, and temporal contexts within which educational activities occur. Observation data can be used to address a wide variety of basic and applied research questions in relevant areas such as learning, instruction, assessment, evaluation, diagnosis, and intervention.

An early proponent of naturalistic observation was Arnold Gesell, a student of founding American psychologist G. Stanley Hall, who developed coding systems to record the behavior of infants during the 1920s. Gesell employed a variety of techniques that would become commonplace in observational studies, including the use of one-way screens to minimize observer influence and motion picture cameras to capture observations for later study. Florence Goodenough developed a short-sample technique, whereby children were observed for brief periods of time and the presence or absence of specific behaviors were noted. Over the next decade, a variety of other psychologists observed children in nursery

schools and other commonplace settings and used coding systems to record social interactions.

Henry Murray employed observational techniques with adults during the late 1930s, which he then applied to personnel selection for the U.S. Office of Strategic Services during World War II. His work included the early use of role-playing as a way to sample behavior during normally occurring social situations. During this same time period, Kurt Lewin favored observation techniques over psychological tests and played a key role in popularizing the use of observational techniques. Roger Barker worked with Lewin in studies of frustration in children, and with colleagues later coded the moment-to-moment behaviors of a child documented in *One Boy's Day*. Another seminal figure in naturalistic observation was Robert Bales, who developed a method for categorizing behaviors documented in the influential book *Interaction Process Analysis.*

The rise of behavior therapy during the 1950s and early 1960s brought observational techniques into widespread use in psychology. Operant behavioral procedures pioneered by psychologist B. F. Skinner and colleagues had particular influence in educational psychology. In early studies, observational measures of child behaviors were employed to monitor the effectiveness of various techniques to change behaviors within the school context. Two key research groups that were among the first to employ behavioral observational techniques extensively were those of Sidney Bijou and Gerald R. Patterson. During the late 1960s, Patterson, John B. Reid, and colleagues spearheaded the development of sophisticated behavioral observation coding systems to study child aggression and other antisocial behaviors within family and classroom settings. Since this period, a wide variety of coding systems have been developed by educational, clinical, and developmental psychologists to monitor various aspects of social interactions within school, home, clinic, and laboratory settings.

The hallmark of naturalistic observation is the accurate recording of specific behaviors at the time they occur within a real-life setting or situation of interest. To accomplish this, either an observer must be present so that coding can be done "live," or an audio or visual recording device must be active so that coding can be done at a later date. To ensure impartiality, observers should not personally know the research subjects or the details of their involvement in the research study, and to ensure accuracy, observers should be adequately trained and monitored in the reliable use of a well-developed and tested descriptive coding system.

Clearly, two important aspects of naturalistic observation are the characteristics of the coding system to be employed and the ability of observers to appropriately apply the coding system. There are two major types of coding systems, those that focus on the discrete behaviors by a given individual, or "micro" coding systems, and those that focus on larger series of a variety of presumably related behaviors by one or more individuals, or "macro" coding systems. Within each type, coding systems have been developed to code a variety of details of social interactions, including physical movement, language content, affective tone, physical location or position relative to others, and the time spent in any of the above.

Regardless of the type or purpose, a coding system must clearly specify what is to be recorded; should contain a limited number of mutually exclusive code categories; and should not require observers to make inferences about phenomena that cannot be seen, such as the internal thoughts or feelings of an individual about a given behavior. Basic training in the reliable use of a comprehensive coding system usually takes several months of intensive work by a coding team. Training in relatively simple coding systems may take less time, but it requires a significant amount of time investment by a team nonetheless.

Once data are being collected for a study, the day-to-day reliability of the observers in continuing to apply the coding system accurately should be monitored. Generally, this is done through the use of regular, random checks where two observers code the same interaction, and their agreement on specific codes is checked. Usually, from 10% to 20% of observations are coded by two or more observers to conduct such reliability checks. Observers who do not meet set reliability standards should be removed from the data collection process and provided further training.

Because of the recognition that human behavior tends to vary across situations, researchers who collect naturalistic observation data often attempt to collect multiple measures of the behaviors that are being observed. This includes multiple observations within the same setting, observations across different settings, and observations as well as impressions by several raters. For example, child physical aggression might be assessed via micro-observational coding of several home, classroom, and playground interactions,

as well as by more global questionnaires completed by parents, teachers, children, and peers. Important questions within most areas where naturalistic observation is used include how many observations and what length of time is required to obtain a truly representative sample of the behavior of an individual, how study participant characteristics such as culture affect the reliability and validity of coding systems, and how the significant costs related to creating and maintaining a reliable coding team can be managed to enable the widespread use of observational techniques in research and practice settings.

J. Mark Eddy

See also Applied Behavior Analysis; Evaluation; Qualitative Research Methods

Further Readings

Bales, R. F. (1951). *Interaction process analysis.* Cambridge, MA: Addison-Wesley.

Barker, R. G., & Wright, H. F. (1951). *One boy's day.* New York: Harper & Row.

Jones, R. R., Reid, J. B., & Patterson, G. R. (1975). Naturalistic observations in clinical assessment. In P. McReynolds (Ed.), *Advances in psychological assessment* (Vol. 3, pp. 42–95). San Francisco: Jossey-Bass.

Mash, E. J., & Terdal, L. G. (Eds.). (2001). *Behavioral assessment of childhood disorders* (3rd ed.). New York: Guilford.

McReynolds, P. (1986). History of assessment in clinical and educational settings. In R. O. Nelson & S. C. Hayes (Eds.), *Conceptual foundations of behavioral assessment* (pp. 42–80). New York: Guilford.

Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. (2004). *Unobtrusive measures* (Rev. ed.). Thousand Oaks, CA: Sage.

NEUROSCIENCE

Neuroscience is the study of nervous system function. It is a discipline that draws on a number of other fields, including anatomy, biology, chemistry, genetics, pharmacology, and psychology. The phenomena studied by neuroscientists range from processes within a cell, including production and trafficking of proteins within cells, to understanding brain systems that underlie consciousness, with studies of changes in cell shape and function, brain chemicals, and hormones as just a few

examples of some of the intermediate areas of study. With regard to education, the brain is the critical substrate that mediates new learning. The brain must be sufficiently flexible to allow new information to be encoded while minimizing previous information that is forgotten. Neuroscience is a critical approach to understanding which situations are likely to maximize new learning. Not surprisingly, with such a complicated subject to study—the nervous system—a number of subdisciplines have developed that offer different approaches within this research area. Important contributions of neuroscience to the study of brain diseases are mentioned in this entry. Some general neuroscience terminology and techniques are described, followed by a section that addresses how neuroscience may be beneficial for improving education.

Cross-Disciplinary Approaches for Studying Brain Diseases

Several diverse approaches for studying the brain have been developed that emphasize different levels of analysis. Molecular and cellular neuroscience emphasizes how the machinery within a cell operates. The results of these research areas have given an appreciation of the complexity of cell structure and function. However, cells within the brain do not act in isolation but rather communicate with each other. Systems neuroscience emphasizes how groups of brain cells interact with other groups of brain cells. This area can yield insights about how one group of brain cells can cooperate (or not) to affect activity in other groups of brain cells. This area has benefited from approaches using computational neuroscience, in which mathematics is used to try to describe how a group of neurons could function to mediate some aspect of processing. Computational neuroscience can yield plausible hypotheses that can then be tested with laboratory-based methodologies.

Neuroanatomical approaches are used to identify specific brain regions and to describe the nature of connections between different brain regions. Understanding how the brain is "wired" is important for testing hypotheses about which brain regions influence each other. Neuropharmacological approaches emphasize studying how drugs affect the brain. Behavioral neuroscientists study the brain mechanisms underlying a range of behaviors, including remembering, feeding, or defensive behaviors. These general research areas are indicative of the diversity of approaches for studying

the brain, but are not by any means exhaustive. Despite increased understanding of brain function using each of these approaches, it has become clear that, to fully understand phenomena related to the brain, multiple levels of analysis need to be implemented. For example, if somebody is interested in the effects of a particular drug, one may study the effects of that drug on the brain itself. Does the drug change the activity of proteins within the cell? Does the drug alter interactions between cells? If so, which cells? Researchers also want to know about the effects of the drug on cognitive processes such as attention and memory as well as changes in behavior. The ultimate goal is to be able to assess exactly which changes in the brain contribute to changes in cognition and behavior. Thus, although there are a number of subdisciplines that emphasize different aspects of brain functioning, a complete understanding of complex phenomena requires cross-fertilization between different approaches within neuroscience.

Neuroscientists are dedicated not only to understanding how the brain works under "normal" conditions, but also to assessing how the brain becomes dysfunctional in some diseases. Several typical strategies are employed in trying to study brain diseases. Of course, it is important to collect information about the patients with the disease to characterize the symptoms and begin to assess the underlying causes. For a particular disease, it is important to characterize the psychological deficits as well as the brain regions, neurotransmitter systems, and molecular markers affected by the disease. Clearly, multiple levels of analysis are required to characterize the disease process.

For many diseases, animal models can be developed, at least for some aspects of the disease, and then changes in the brain's neurochemistry and molecular and cellular functioning can be assessed. Animal models also offer an opportunity to test preliminary treatments for some brain diseases. An important consideration in animal models is the extent to which findings translate to the human condition. Neuroscientists have an excellent history of successfully translating research from animal models to the human condition. For example, early animal research using the drug reserpine, which depletes certain chemicals in the brain, was observed to produce symptoms similar to those in depression. Many drugs now, including a class of drugs known as serotonin-specific reuptake inhibitors, such as Prozac, were developed based on the early findings with animals. Another

example involves the treatment of Parkinson's disease. Animals administered a particular drug, MPTP, exhibit symptoms and brain pathology similar to patients with Parkinson's disease. Based on studying the brain changes in these animal models, damage to particular brain structures (and later deep brain stimulation of these structures, which is reversible) was shown to alleviate some symptoms in Parkinson's disease. These examples demonstrate how basic neuroscience research can contribute to treatments for devastating neuropsychiatric diseases. Future success in treating brain-based diseases likely depends on integration of research from multiple levels of analysis, from molecular/cellular to behavior.

General Terminology in Neuroscience

This section initially describes concepts at the level of brain cells and then moves to a discussion of brain regions. The cells responsible for much of the communication within the brain are neurons. Signaling between neurons is accomplished through both electrical and chemical means. One of the major components of neurons is the cell body, which is where the nucleus containing genetic information is located along with a number of proteins necessary for maintaining the neuron's health. A thin protrusion from the cell body, the axon, carries electrical signals to the axon terminal. The axon terminal contains chemicals used in neuronal communication, known as neurotransmitters. When an electrical signal (an "action potential") is conducted along an axon and reaches the axon terminal, the terminal will release neurotransmitters into an area known as the synaptic cleft or synapse. The neuron-releasing neurotransmitter can be identified as the presynaptic neuron. The neurotransmitter binds to receptors on a nearby neuron (the postsynaptic neuron) and has effects on that postsynaptic neuron. These effects include increasing or decreasing the likelihood that the postsynaptic neuron will fire an action potential or altering the rate at which proteins are synthesized. The actions of the neurotransmitter can be terminated in several ways, including by enzymes or by being taken back up into the presynaptic neuron.

Many different neurotransmitters have been identified, and one significant goal of neuroscience is to understand what these neurotransmitters uniquely contribute to neuronal processing as well as how these

neurotransmitters become altered in some neuropsychiatric diseases. Some of the most prevalent neurotransmitters in the brain are amino acids, including glutamate and gamma-aminobutyric acid (GABA). Among other functions, glutamate is thought to be an important neurotransmitter for learning and memory. GABA is implicated in anxiety. Drugs such as Valium, which are used to treat anxiety, act at the same receptor complex as GABA.

Other important neurotransmitters include a group of neurotransmitters known as catecholamines, which include norepinephrine and dopamine. Norepinephrine is involved in the arousal of an organism. Increases in dopamine in some brain regions are a consistent feature of most drugs of abuse. Furthermore, increased activity of the dopamine system is associated with symptoms in schizophrenia. Damage to a particular dopaminergic pathway appears to underlie some of the motoric deficits in Parkinson's disease. Another neurotransmitter, serotonin, has been implicated in impulsivity and depression. Some drugs used to treat depression, such as Prozac, allow serotonin to act on receptors for a longer period of time. Another important neurotransmitter is acetylcholine. Loss of neurons that release acetylcholine has been associated with the cognitive deficits in Alzheimer's disease. Many of the drug treatments for Alzheimer's disease act by blocking the breakdown of acetylcholine. The transmitters mentioned above are typically considered "classical" neurotransmitters and have been studied extensively. However, numerous other molecules, even some gases, appear to have properties similar to classical neurotransmitters and will be important for neuroscientists to continue to study in the future.

The actions of neurotransmitters are mediated through receptors. Receptors are simply proteins that commonly span the outside and inside of the cell, with most neurotransmitters binding to a portion of receptors outside of the cell. For each of the classical neurotransmitters, there are a number of receptors. Some goals of neuropharmacology include identifying the different types of receptors for a particular neurotransmitter as well as where these receptors are located within the brain. This research is important because some neurotransmitters appear to have different actions mediated through different receptor subtypes. In some cases, it is desirable to develop a drug that acts at only a particular receptor subtype in order to limit the drug's side effects. For example, drugs used to treat depression by increasing serotonin levels appear to provide most of the antidepressant effects by acting at the serotonin-1 receptor. However, increasing serotonin levels also causes more serotonin to bind to other serotonin receptors, including serotonin-2 and serotonin-3 receptors. Increased serotonin binding at serotonin-2 and serotonin-3 receptors appears to contribute to some side effects of this class of antidepressants, including sexual dysfunction and nausea. Thus, a goal of researchers is to develop drugs that target the receptor that provides the antidepressant effects, but not the receptors responsible for undesirable side effects.

Another level of analysis regarding brain function is to study processing mediated by particular brain regions. The nervous system as a whole can be divided into two systems, the peripheral nervous system and the central nervous system. The peripheral nervous system conveys information to muscles, sensory receptors, and many organs. The peripheral nervous system comprises two major systems—the somatic nervous system, which largely is responsible for controlling voluntary movements, and the autonomic nervous system, which regulates many involuntary processes, such as heart rate and breathing. The autonomic nervous system comprises two systems, the parasympathetic nervous system and the sympathetic nervous system. The parasympathetic nervous system is typically active when we are attempting to increase the amount of energy stored. Thus, parasympathetic nervous system activation can stimulate digestion and salivation. The sympathetic nervous system is critical when we need to expend energy and is often referred to as being important for our fight-or-flight response, and thus performs functions such as increasing heart rate to increase blood flow to muscles.

The central nervous system comprises the spinal cord and the brain. The brain can be divided into four lobes:

1. The occipital lobe, which is important for processing visual information

2. The temporal lobe, which is important for processing auditory information and for some memory functions

3. The parietal lobe, which is involved in processing some spatial information and for integrating information from other brain areas

4. The frontal lobe, which is involved in processes including planning, working memory, and selectively attending to objects in our environment

The outer 2–4 mm of the brain is grey matter, consisting mainly of cell bodies, known as the cortex.

A number of cortical regions have been identified, including cortical areas that process sensory information, such as for vision, audition, or olfaction. There are other cortical areas (and brain regions outside of the cortex) that engage our higher cognitive processes, such as planning, selectively attending to stimuli in the environment, or storing information in memory. Below the cortex is white matter, which primarily includes axons projecting to particular cortical regions. Additionally, many subcortical brain regions, which are structures located below the cortex, are present. These structures include the thalamus, which, in part, relays information to cortical regions, and the hypothalamus, which maintains an organism's homeostasis. There are brain regions toward the back of the head that contribute to balance (the cerebellum) and that maintain many life-sustaining processes (the medulla).

Although some specific processes can be associated with some brain regions, contemporary neuroscience has demonstrated the importance of studying how brain systems, that is, interconnected brain regions, contribute to complex phenomena. For example, if you are at a stoplight, you must process the location and color of the light (as well as attend to the stoplight). Most likely, visual cortical brain regions will be important for this processing. However, you must also engage the rules for changing your behavior—take your foot off the brake and press the accelerator when the light changes from red to green. This engagement of rules likely involves several cortical (and subcortical) regions, including the prefrontal cortex. The primary motor cortex or subcortical motor structures may be activated to cause your foot to move to the appropriate location as well as to receive proprioceptive feedback concerning how hard you are pressing. This example concerns a behavior that most drivers take for granted, yet gives at least some superficial indication of how several brain structures must work in a coordinated fashion to complete this task.

Techniques Used in Neuroscience

One of the most critical aspects to advancing our understanding of the brain is the development of tools to manipulate and measure brain functioning. This section highlights a few of these techniques, with an emphasis on comparing the relative advantages and disadvantages of different techniques. In humans, a number of noninvasive procedures can be employed to study neural functioning. One technique that has

become prominently used is called *functional magnetic resonance imaging*. This technique provides a measure of changes in blood oxygen levels to particular brain regions. Increases in blood oxygen to a brain region are thought to reflect increased activity by that brain region. Magnetic resonance imaging offers impressive spatial resolution, that is, one is able to localize within the brain where the increase in blood oxygen flow occurs. The time course of this change in blood flow is at least a couple of seconds. In some instances, this temporal resolution may not be sufficient, and other techniques may need to be employed that allow measures to be taken more quickly. One approach for using this technique is to manipulate cognitive demands. For example, participants could be asked to perform a working memory task or a similar task without a working memory requirement. One could then assess which brain regions show increases in blood oxygen levels during the working memory task, but not during the task with the working memory requirement removed. Ultimately, these data could provide information about the neural systems that are selectively activated during working memory.

A procedure for measuring electrical activity from the scalp is electroencephalography. With this technique, a cap containing electrodes is placed over specific regions of the head. This procedure can measure the frequency of brain waves as well as changes in brain wave activity (termed *event-related potentials*) in response to a stimulus. Electroencephalography offers temporal resolution on a millisecond scale, but it is difficult to localize the origins of the electrical signal. Thus, functional magnetic resonance imaging and electroencephalography differ in temporal and spatial resolution, leaving it to the researcher to determine which technique is more appropriate for a particular experiment. For more invasive procedures, humans can be injected with trace levels of radioactive ligands that bind to particular receptors in the brain. The amount of radioactivity can then be measured, typically using a procedure known as positron emission tomography, to determine the number and efficacy of a particular receptor population.

Animals can be used to test hypotheses about brain function. One traditional approach with animals is to damage a portion of the brain (or a particular population of neurons within a brain region) and assess changes in behavior. Lesion techniques can be used to understand whether a brain region is necessary for

some behavior or function. Thus, lesion techniques allow manipulation of the brain and assess changes in behavior or other aspects of brain functioning. However, much like with humans, in some cases it is also desirable to assess the effects of manipulating behavior or cognitive processing demands (e.g., compare performance on a working memory task with a procedure that does not tax working memory) on brain function. Electrophysiology allows measures of the rate of action potentials from a neuron or small groups of neurons. Electrophysiology provides excellent temporal resolution, but it may be difficult to assess which neurotransmitter systems are leading to changes in neuronal firing rates.

Microdialysis is a procedure in which a probe with a semipermeable tip is inserted into a brain region and artificial cerebrospinal fluid is perfused through that brain region. Neurotransmitters can diffuse through the semipermeable membrane and be collected, and levels can be assessed using a procedure known as high-performance liquid chromatography. Microdialysis does allow measurement of specific neurotransmitters, but often samples must be collected over at least 1 minute. Thus, microdialysis has a poorer temporal resolution than electrophysiological techniques. Ultimately, it is critical for neuroscientists to ascertain the relative strengths and limitations of the techniques employed in an experiment. Such information is critical for developing a clear interpretation about the effects of a manipulation on brain functioning.

A number of techniques have been developed to study how genetics, including the expression of genes, contribute to brain function. Genes provide a code for specific proteins, and the regulation of these proteins can affect brain functioning. For example, genes can code for particular portions (called subunits) of a receptor. The type of subunit that is coded for can affect the overall activity of the neurotransmitter system that acts at that receptor. One approach, gene microarrays, allows researchers to examine, from properly prepared brain tissue, the expression of thousands of genes from that tissue. Although microarrays are a powerful approach, it is possible that, with such a large number of genes being studied, some genes will demonstrate increased or decreased expression just by chance.

Microarray analyses can provide a set of genes for future study, but additional approaches are important to further characterize the role of the expression of a specific gene in brain function. More targeted approaches include looking for a single change in the genetic code, known as a single nucleotide polymorphism. The study of the role of gene expression in neuroscience has accelerated over the past 10 years, and techniques that assess many genes at once, as well as more targeted approaches that study a single gene, will be important in providing evidence to further the study of brain function.

One goal within neuroscience is to understand relationships between brain and behavior or brain and underlying cognitive processes. To accomplish this goal, it is essential to have behavioral paradigms that assess the processes under study, in addition to sophisticated measures and manipulations of the brain. In humans, computer-controlled stimulus presentation can be beneficial for standardizing experimental variables across laboratories as well as for carefully timed response measures. Numerous paradigms with animals have been developed to assess (among other processes) attention, working memory, anxiety-like behavior, and depressive-like behaviors. Some of these paradigms offer "artificial" laboratory settings in which animals are rewarded for pressing a lever in response to a particular stimulus. Similar to human research with computers, these laboratory-based tasks offer consistency in experimental procedures across laboratories, thus facilitating comparisons about research conducted in different laboratory environments. Typically, there are procedures employed to validate that a task is assessing the desired construct. For example, in a working memory task, performance would be expected to decrease as participants must remember information for a longer time period or when the number of items to be remembered is increased. The continued development of behavioral paradigms that allow specific assessment of behavior and/or cognitive processes is essential for neuroscience research aimed at understanding brain-behavior relationships.

Neuroscience and Education

How can neuroscience research contribute to education? One major goal of education, of course, is to increase knowledge and critical thinking. Behavioral neuroscience is a field dedicated to studying relationships between brain and behavior, including neural mechanisms underlying learning and memory. By understanding how the brain processes and stores information, educational programs can be developed that will provide material in a way that increases the

likelihood that the brain will be able to properly recognize and recall that material.

Neuroscience research can also help us to understand the development of brain regions and how that development may affect learning. For example, if we know that the prefrontal cortex develops later in adolescence, education programs that do not place heavy demands on prefrontal cortical processing may be beneficial. Conversely, it is also possible that educational programs that require prefrontal cortical processing may accelerate the development of this brain region. It is clear that the brain is a plastic organ that is continuously changing in terms of the strength of connections between different neurons. Education programs and neuroscience research could work together to learn more about strategies that stimulate brain development and about educational programs that may be more effective at different stages of brain development.

Developmental disorders can represent a significant challenge to successful education. One of the most common examples of such a disorder is attention deficit hyperactivity disorder (ADHD). Neuroscience research focuses on the changes in the brain using noninvasive techniques in humans or animal models. For example, individuals diagnosed with ADHD may perform a series of tasks while functional magnetic resonance imaging is employed to assess whether specific brain regions show abnormally high or low blood flow levels compared with a control group. Based on animal models or positron emission tomography, in which individuals are injected with trace levels of radioactively labeled ligands, it is possible to assess whether particular neurotransmitter levels are elevated or decreased in individuals with ADHD. By using this approach, a pharmacological strategy can be developed to regulate neurotransmitter levels. Thus, neuroscience research can be used to try to improve traditional education programs and to treat some neuropsychological disorders that can impede classroom learning.

One of the major challenges to education is that important social development is occurring that may assist or hinder learning. Furthermore, the educator may benefit from information about what types of material is better learned individually versus in a group setting. Traditionally, neuroscience has not focused heavily on social behavior; however, new fields are emerging, such as social neuroscience, that will provide information about how the brain responds to social interactions. A more complete understanding of the neural basis of social interactions may allow educators to predict what types of materials may be learned well in a group setting as opposed to course information that may benefit from more individualized studying. At the same time, for educators to receive maximum benefits from the tools and expertise of neuroscientists, it is essential that educators describe what types of information would be most useful to enhance student performance.

Considerations

Neuroscience is a discipline that ranges in study from the cells, neurons, which are responsible for much of the communication within the brain, to assessing the cognitive and behavioral consequences of brain damage. Given the range of approaches necessary for such diverse topics, it is not surprising that several subdisciplines have developed within neuroscience. Furthermore, the techniques necessary for breakthroughs in these subdisciplines are diverse, and the exact nature of what they are measuring needs to be considered carefully. Another important consideration is that the different subdisciplines inform each other in order to characterize more fully brain functioning. Such cross-disciplinary approaches will provide more information about the neural mechanisms related to learning as well as for treating neuropsychological disorders that can disrupt learning.

Joshua A. Burk

See also Brain-Relevant Education; Dynamical Systems

Further Readings

Bechtel, W. (2002). Aligning multiple research techniques in cognitive neuroscience. *Philosophy of Science*, *69*, S48–S59.

Cooper, J. R., Bloom, F. E., & Roth, R. H. (2003). *The biochemical basis of neuropharmacology* (8th ed.). New York: Oxford University Press.

Posner, M. I., & DiGirolamo, G. J. (2000). Cognitive neuroscience: Origins and promise. *Psychological Bulletin*, *126*, 873–889.

Society for Neuroscience. (2003). *Translational neuroscience accomplishments.* Retrieved from http://www.sfn.org/skins/main/pdf/gpa/translational.pdf

Squire, L. R., Bloom, F. J., McConnell, S. K., Roberts, J. L., Spitzer, N. C., & Zigmond, M. (2003). *Fundamental neuroscience* (2nd ed.). San Diego, CA: Elsevier Science.

No Child Left Behind

On January 8, 2002, President George W. Bush signed the No Child Left Behind (NCLB) Act into law. No Child Left Behind was a comprehensive, complex, and controversial education law that was passed as a reaction to the low academic achievement, especially reading achievement, exhibited by many public school students in the United States. The law represented an unprecedented increase in the role that the federal government plays in education because the law dramatically increased federal mandates and requirements on states, school districts, and public schools. The NCLB also increased federal funding to states by almost 25%. The law affected all students who attended public schools by requiring states and public schools to be held accountable for improving student achievement in reading and math. Moreover, NCLB required that educators use scientifically based strategies and methods, which represent the primary tools that will allow schools to make meaningful changes in the academic achievement of their students. This entry begins with a brief description of the history of NCLB and describes the major components of the law.

History

No Child Left Behind was the most recent reauthorization of the Elementary and Secondary Education Act (ESEA) of 1965. The ESEA was passed as part of President Lyndon Johnson's War on Poverty. The law appropriated federal money for states to improve educational opportunities for disadvantaged children. Although Congress allocates funds to the ESEA annually, Congress must reauthorize the law every 5 or 6 years.

In 1994, the bill that reauthorized the ESEA was titled the Improving America's Schools Act (IASA). The central purpose of IASA was to implement standards-based education throughout the nation. The IASA created a new framework for the federal role in elementary and secondary education, in which the federal government not only provided aid to schools serving economically disadvantaged students but extended federal support to the states' implementation of local and state standards-based reform. The IASA was based on states' developing challenging academic standards, creating and aligning assessments for all

students, holding schools accountable for results, and increasing aid to high-poverty schools.

The 2001 reauthorization of the ESEA was titled NCLB. The passage of NCLB expanded the role of the federal government in public education by holding states, school districts, and schools accountable for producing measurable gains in students' achievement in reading and mathematics. For the first time, the federal government in NCLB began requiring states and school districts to use numerical data to provide evidence of improved student outcomes.

Purpose

The primary purpose of NCLB was to ensure that students in every public school achieve important learning goals while being educated in safe classrooms by well-prepared teachers. To increase student achievement, the law required that school districts (a) assume responsibility for all their students reaching certain goals in reading and math, (b) use scientifically based procedures to teach reading and math, and (c) have highly qualified teachers in all their classrooms. Furthermore, NCLB required schools to close academic gaps between economically advantaged students and students who are from different economic, racial, and ethnic backgrounds as well as students with disabilities. The primary goals of NCLB were that

- All students will achieve high academic standards, by attaining proficiency or better, in reading and mathematics by the 2013–2014 school year.
- Highly qualified teachers will teach all students by the 2005–2006 school year.
- All students will be educated in schools and classrooms that are safe, drug free, and conducive to learning.
- All limited-English-proficient students will become proficient in English.
- All students will graduate from high school.

Major Principles

Accountability for Results

No Child Left Behind focused on accountability for results in three major ways. First, the law required states to first identify the most important academic content for students to learn. Second, NCLB required that states adopt or develop statewide assessments that were aligned to the state standards in reading-language arts, math, and

eventually science. Third, states were required to set proficiency standards, which were goals that schools and school districts had to attain within certain periods of time in order to have 100% of their public school students scoring proficient or above on the statewide tests of reading and math by the 2013–2014 school year.

Statewide Academic Achievement Standards. No Child Left Behind required states to identify important academic content for students to learn. Specifically, the law required states to develop academic standards for all students in reading-language arts, math, and science. States were free to develop standards in other areas, too. The purpose of the state-defined standards was to provide guidelines to schools, parents, and teachers that tell them what achievement will be expected of all students.

Statewide Assessments. No Child Left Behind required that states adopt or develop statewide assessments that were aligned to the state standards in reading-language arts, math, and eventually science. The purpose of the statewide testing was to measure how successfully students were learning what was expected of them and how they were progressing toward meeting these important academic standards. States were required to assess students' knowledge and skills in reading and math in Grades 3 through 8. The purpose of the state tests was to enable stakeholders (e.g., teachers, administrators, parents, policymakers, and the general public) to understand and compare the performances of schools against the standards for proficiency as set by the states. The results of these assessments were to be reported to parents in annual report cards. This information would tell parents about where their child stood academically and if their child's school and school district were succeeding in meeting state standards. Thus, these assessments were used to hold schools accountable for the achievement of all students.

Adequate Yearly Progress. States were required to set proficiency standards, which were goals that schools and school districts had to attain within certain periods of time in order to have 100% of their public school students scoring proficient or above on the statewide tests of reading and math by the 2013–2014 school year. In addition to all students in a school, schools also were to test and report on the performance of the following subgroups: students who were economically disadvantaged, students from racial and ethnic groups, students with disabilities, and students with limited

English proficiency. These goals or targets represented percentages of students who achieved proficiency on the statewide tests in reading and mathematics. Schools had to make these targets in order to achieve adequate yearly progress (AYP).

States were responsible for determining their own system of requirements and rewards to hold all public schools and school districts responsible for meeting AYP. For example, if a school met its AYP target, it could be designated a "Distinguished School." If a school failed to meet its AYP target, the school was to receive assistance from the state to improve its scores. When a school was first identified for improvement, the state provided technical assistance to enable the school to address the specific problems that led to its being identified. The school, in conjunction with parents and outside experts, had to develop a 2-year improvement plan.

No Child Left Behind also had very specific requirements for schools that did not make AYP. If a school did not make AYP, this information had to be published and disseminated to parents, teachers, and the community in an easy-to-understand format. Schools that did not make AYP for 2 consecutive years were designated in need of improvement. In such situations, the state needed to continue to provide technical assistance to the school. Additionally, the school had to offer the parents of students in the designated school the option of transferring to another public school within the district. This option was called public school choice. For schools that continued to fail to make AYP, NCLB required that schools take the following actions:

1. *Three consecutive years:* In addition to continuing to offer public school choice, the school district had to offer supplemental services to disadvantaged students.

2. *Four consecutive years:* In addition to continuing to offer public school choice and supplemental education services, the school was required to implement corrective actions to improve the school, such as
   - replacing certain staff responsible for failure to make AYP,
   - implementing a new curriculum grounded in scientifically based research,
   - hiring outside experts to assist the school, and
   - reorganizing the management structure.

3. *Five consecutive years:* In addition to the above, the school had to be restructured by taking actions such as

- replacing the staff,
- contracting with a private firm to run the school, and
- reopening the school as a charter school. If these procedures are not successful, the state may take over management of the school district.

Schools that are in need of improvement can be removed from that category if the school makes AYP for 2 consecutive years.

Scientifically Based Instruction

The second major principle of NCLB required that states and school districts use scientifically based instructional programming to improve the achievement of students. The authors of NCLB believed that, too often, schools had adopted programs and practices that were not effective, which would result in lower academic achievement. No Child Left Behind emphasized using educational programs and practices that had been demonstrated to be effective by rigorous scientific research. A central principle in NCLB, therefore, required that federal funds be expended to support only educational activities that were supported by scientifically based research evidence that the procedures actually increased student achievement.

The intent of NCLB is to require that rigorous standards be applied to educational research and that research-based instruction is used in classroom settings. This means that state and local educational agencies must pay attention to research in education and ensure that teachers use evidence-supported methods in their classrooms.

The National Research Council reported that for a research design to be scientific, it must allow for direct, experimental investigation of important educational questions. No Child Left Behind defined scientifically based research as "research that applies rigorous, systematic, and objective procedures to obtain relevant knowledge" (NCLB § 1208(6)). This includes research that (a) uses systematic, empirical methods that draw on observation or experiment; (b) involves rigorous data analyses that are adequate to state hypotheses and justify the conclusions; (c) relies on measurement or observational methods that provide valid data for evaluators and observers and across multiple measures and observations; and (d) has been accepted by a peer-reviewed journal or approved by a panel of independent experts through a comparably rigorous, objective, and scientific review.

Highly Qualified Teachers

The authors of NCLB believed that the quality and skill of a student's teacher are extremely important factors in student achievement. Congress recognized the importance of having well-prepared teachers in public school classrooms when they included provisions in the NCLB requiring that all new teachers hired in programs supported by Title 1 funds had to be highly qualified teachers beginning with the 2002–2003 school year. Additionally, the law required that by the end of the 2005–2006 school year, all teachers in public schools had to be highly qualified. The NCLB also requires that states ensure that paraprofessionals who work in the nation's classrooms must also be highly qualified.

There are three basic requirements in the NCLB that public school teachers must meet to be highly qualified. First, teachers must hold a minimum of a bachelor's degree from a college or university. Second, teachers must have full state teacher certification or licensure for the area in which they teach. Third, teachers must be able to demonstrate subject matter competency in the core academic subjects in which they teach. Teachers can demonstrate subject matter competence by passing a state-administered test in each of the core subjects that he or she teaches. The structure and content of these tests are determined by the individual states.

To ensure that only highly qualified teachers teach in public school classrooms, each state receiving funds under Title 1 of NCLB must develop a plan to ensure that all of the state's public school teachers are highly qualified to teach core academic subjects in which they provide instruction. The NCLB regulations defined core academic subjects as English, reading-language arts, mathematics, science, foreign languages, civics, government, economics, art, history, and geography. If a teacher taught in one of these core subjects, the NCLB highly qualified requirement applied to him or her. If a teacher taught in more than two of these core subjects, he or she had to be qualified in all of the subject areas taught.

In summary, NCLB was a complex, sweeping, and controversial law that was passed as a reaction to the low academic achievement exhibited by many public school students in America. NCLB held states, school districts, principals, and teachers accountable

for making meaningful improvements in students' academic performance. No Child Left Behind also points educators toward the tool that will allow schools to make meaningful changes in the academic achievement of their students: scientifically based research.

Mitchell Yell

See also Individual Differences; Individuals with Disabilities Education Act; Learning Style

Further Readings

Caterino, L. C. (2007). Review of assessment accommodations for diverse learners. *Journal of Psychoeducational Assessment, 25,* 206–208.

Charles, M. T. (2004). *Where are we going as we leave no child behind? La technique and Postman, Papert, and Palmer—Part One by Michael T. Charles.* Retrieved July 10, 2007, from http://bcis.pacificu.edu/journal/2004/01/charles.php

The White House. (2001). No Child Left Behind. Retrieved July 10, 2007, from https://www.whitehouse.gov/news/reports/no-child-left-behind.pdf

NORMAL CURVE

The *normal distribution* is one of the most important probability distributions in statistics in that many statistical analyses build on the assumption that the data follow the normal distribution, and in that many physical and biological phenomena in real life can be approximated by the normal distribution.

A normal distribution is specified by two parameters: mean μ and standard deviation σ. If a random variable X follows the normal distribution with mean μ and standard deviation σ, it is often denoted by $X \sim N(\mu, \sigma^2)$. The normal distribution has a bell shape as shown below, which is called the normal curve.

Technically, the normal curve is given by the following formula:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right].$$

The variable x takes any real value. The mean μ specifies the central location of the distribution, and the standard deviation σ determines to what extent the distribution is spread out. The curve has a maximum
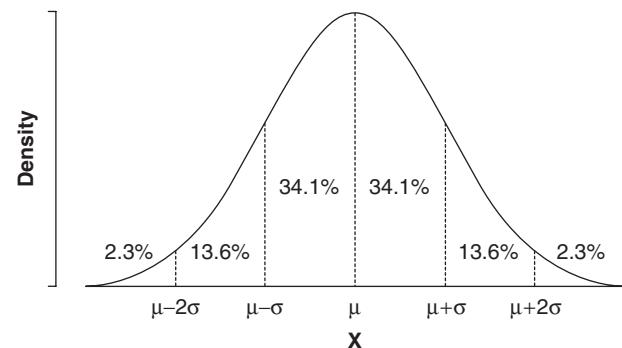
Figure 1   The Normal Curve

height at $x = \mu$ and is symmetric about μ. In probabilistic terms, the normal curve represents the *probability density function* of the normal distribution; the height of the curve at each point of x denotes the corresponding probability density. The most important characteristic of a probability density function is that an area under the curve on a specified interval taken on the x-axis represents a probability that the random variable takes values within that interval. In the case of the normal distribution, as shown in Figure 1, the probability that X is within one standard deviation from the mean is approximately 68.2%, and the probability that X is within two standard deviations from the mean is approximately 95.4%.

Especially if $\mu = 0$ and $\sigma = 1$, the distribution is called the *standard normal distribution*, which is frequently used as the reference distribution in statistical testing. A probability table for the standard normal distribution is usually provided in standard statistics textbooks. Any random variable that follows an arbitrary normal distribution can be transformed to the standard normal by subtracting the mean and then dividing by the standard deviation. In other words, if $X \sim N(\mu, \sigma^2)$, then $(X-\mu)/\sigma \sim N(0, 1)$. Thus, probability values for any normal distribution can be referred to in a probability table for the standard normal distribution.

The *central limit theorem* exemplifies the theoretical importance of the normal distribution. Suppose that you take a random sample of size n from an arbitrary distribution and compute the mean (or sum) of the sampled values. The central limit theorem states that the distribution of the sample mean (or sum) tends to be approximated by a normal distribution as the sample size n increases, no matter what the

original distribution is, as long as it has finite mean and variance. For example, when you flip a coin n times and count the number of heads, the resulting distribution is the binomial distribution with probability .5 and the number of trials n. As n increases, however, the binomial distribution becomes closer to the normal distribution with mean $n/2$ and standard deviation $\sqrt{n}/2$.

The normal distribution serves as an assumption for data distribution in many standard statistical analyses, including the two-sample t-test, the paired t-test, analysis of variance, linear regression, and so on. It also serves as a theoretical foundation for deriving other probability distributions, such as the t distribution, the chi-square distribution, and so on.

The normal distribution was first formulated by Carl Friedrich Gauss in 1811, and was thus given another name, the *Gaussian distribution.* Gauss derived the normal distribution as a distribution such that the sample mean equals the maximum likelihood estimator of the population mean when he was seeking a probability distribution of errors; that is, the difference between a measurement and a true value. Thus, the normal distribution is also regarded as the *law of errors.* Even earlier, Abraham de Moivre showed in 1733 that the binomial distribution is well approximated by the normal distribution as the number of trials increases. In the early 19th century, Pierre-Simon Laplace elaborated de Moivre's finding and restated it in more formal and general terms in mathematical statistics.

Wiliam M. Bart and Kentaro Kato

See also Descriptive Statistics; Inferential Statistics; Standard Scores

Further Readings

Read, C. B. (1985). Normal distribution. In N. L. Johnson & S. Kotz (Eds.), *Encyclopedia of statistical sciences* (Vol. 6, pp. 342–359). New York: Wiley.

Witte, R. S., & Witte, J. S. (2006). *Statistics* (8th ed.). New York: Wiley.

Norm-Referenced Tests

A norm-referenced test is one that is designed to facilitate interpretations of scores by comparing scores based on the ordering of examinees within a well-defined group of interest. Two important definitions associated with norm-referenced interpretations are *percentiles* and *percentile ranks.* A percentile is a test score below which falls a certain percentage of the scores. A percentile rank is the percentage of people who have a score lower than the score of interest.

Since about 1905, norm-referenced interpretations have been a dominant approach to making test scores meaningful to both educators and the public, although criterion-referenced interpretations have been gaining in prominence over the past several decades. A special type of criterion-referenced interpretation, known as *standards-based interpretation*, has been gaining in popularity since No Child Left Behind was enacted into federal law in 2002.

Score Interpretation

By itself, a raw score on a test has no meaning. Knowing an examinee answered 21 questions correctly on a math test is useless by itself. Faced with such useless information, one might ask any of a number of questions.

- How many items were on the test?
- What content within math was covered?
- What related content was excluded?
- What was the cognitive complexity of those items?
- What item formats were used (multiple-choice, problem solving, proofs)?
- How well did other examinees perform on this same test?
- What do we know about other achievements of examinees with similar scores?

Answers to each of these questions will raise other questions. Over time, the meaning of test scores accrues as users become familiar with the characteristics of those scores and the relationships those scores have with variables of interest.

Test makers try to facilitate this development of meaning by creating score scales that support the intended primary inferences. One such approach is referred to as *norm-referenced*—the comparison of the performance of an examinee with the performance of other examinees in a meaningfully defined group. Such interpretations may be particularly useful when determining how to allocate insufficient resources, such as if there are more applicants for an educational

program, school, university, or job than there are openings. For example, norm-referenced tests might be used as a significant piece of information in determining which students should be placed in a remedial or gifted program.

If resource allocation decisions were simple and there were 12 spots in a remedial program, one could simply admit the 12 students with the lowest scores. But most real-world resource allocation problems are more complex and somewhat elastic. Thus, developing expectations over time (and thus sometimes admitting more or fewer students into such a program) is facilitated by normative data. Therefore, policymakers might often prefer for a program to be made available for any students in the bottom 10%, rather than for a fixed number of students.

Normative expectations can also serve to facilitate group comparisons, for example, whether or not students in a school or district are performing as a group better than those in other schools or districts. Whether or not such differences are interpreted correctly, they can influence the perceived desirability of neighborhoods and, thus, real estate prices.

Historical Roots

The use of normative interpretation for test data has its roots in the work of early psychologists such as Wilhelm Wundt and Francis Galton in the late 1800s. These psychologists looked at the distributions of various measures and noted the typical normal distributions.

In 1905, French educator Alfred Binet invented the intelligence test. Scores were expressed as a mental age that could be compared with a student's chronological age to help make educational placement decisions. In 1916, U.S. psychologist Lewis Terman developed a revised version of Binet's test, the Stanford–Binet, and changed the score-reporting scale to the ratio of mental age to chronological age. The resulting IQ scores could be compared regardless of the age of the students, but proved to have more high and low scores than the normal distribution predicted. For this and other reasons, the ratio IQ score was replaced by a deviation IQ score with a mean of 100 and a standard deviation of either 15 (Wechsler) or 16 (Stanford–Binet). Deviation IQ scores were supplemented with information regarding the percentage of people with lower scores. In 1963, Robert Ebel coined the terms *norm-referenced* and *criterion-referenced* tests.

Normative Approaches

A common approach to providing normative information is the use of grade- or age-equivalent scores, just as Binet used more than 100 years ago. The general public finds this approach intuitively appealing because it attaches test performance to a concept with which they are very familiar—grade (or age) level. Student performance is indicated as being at the grade and month where the average child receives that score. Operationally, this is done by creating a concordance table, listing every possible score and the corresponding grade and month for which that score is the average score. Following is a raw score—grade-equivalent score concordance table for a hypothetical fourth-grade reading test with 15 possible raw scores (0–14).

There are two issues in particular that influence the interpretability of grade- and age-equivalent scores. First, such scales are not interval level. That is, the difference between a reading score of 1.1 (a student in the first month of the first grade) and 2.1 represents a greater difference in achievement than the difference between a 7.1 and an 8.1. Thus, it can be misleading to report growth scores or averages.

Second, consider two students who score a 6.5 grade-equivalent score on a science test. One student is in the second month of third grade and the other is in the second month of eighth grade. These students have been exposed to different parts of the science curriculum. It is likely that the younger student answered correctly almost all questions about the parts of the curriculum to which she or he had been exposed, but did not do as well on questions from those areas of the curriculum not yet studied. On the other hand, the older student may well have answered a small portion of the items correctly for all topics on the test. Furthermore, we would expect that 3 years from now, the younger student will surpass the science achievement of the older student.

An alternative approach to providing normative information is to provide percentile ranks for each raw (or scaled) score. This focuses comparisons within a grade (or age) level and thus avoids the second issue mentioned for grade-equivalent scores. The first issue (non-interval scale measurement) is also true for percentile ranks. On the other hand, the general public is not as familiar with percentile ranks as it is with grade level or age. An additional issue with percentile ranks is that they make it difficult to show

Table 1   Hypothetical Raw Score–Grade-Equivalent Score Concordance Table

Raw Score	Grade-Equivalent Score	Raw Score	Grade-Equivalent Score	Raw Score	Grade-Equivalent Score
0	2.2	5	3.9	10	4.7
1	2.6	6	4.0	11	4.9
2	2.9	7	4.2	12	5.2
3	3.2	8	4.4	13	5.4
4	3.6	9	4.5	14	5.7

growth. If students whose true achievement is at the 70th percentile in the fall make typical progress during the school year, they will have higher raw scores in the spring, but they will remain at the 70th percentile. Thus, within-year (or across-year) typical growth is reflected by no change in percentile ranks. Smaller-than-typical growth will be reflected by a drop in percentile rank.

Norming Process

The development of stable, accurate norms is a multifaceted, complex, and logistically difficult process. First, one must identify the normative population. Is it *all* ninth-grade students, just public school students, or public school students for whom English is their primary language of instruction? Should special education students be included? What about home-schooled children? At its best, the normative population should reflect the group within which students will be compared. However, sometimes there are multiple comparison groups, and so multiple norms can be desirable.

Once the normative population is defined, a sampling plan must be developed. Some tests (such as the SAT or ACT) provide normative information based on naturally occurring examinees. Such norms reflect the user population, but are inappropriate or misleading if one wanted to make inferences that went beyond that population, such as all college-bound seniors rather than college-bound seniors who took the SAT (or ACT).

Often, it is considered desirable to define the population of interest and, in a special study, collect data from a national probability sample (a sample that contains randomly selected examinees). Improved norms

estimates (greater precision) can be derived from stratified random sampling; that is, breaking down the sample into subsamples. Stratified random sampling increases the precision of norms estimates most when the stratification variable is correlated with test scores.

Normative data must be gathered at a point in time for which the norms are most appropriate, but different test users will choose to use the test at different times. It is too difficult and expensive to collect normative data for each month or week of the year, so norms are typically collected at two different times for each grade (or age), and results are interpolated or extrapolated for other months.

Once raw data are received, those data must be adjusted. Data from each stratum are weighted to account for differences between the actual proportions collected in each stratum and the true proportions of that stratum in the population. For example, for the norming of an achievement test series, if a researcher stratified based on public and private schools, and after collecting the data, 70% came from public school and 30% from private schools, those data must be weighted to reflect that, in actuality, 86% of students attend public school and only 14% attend private schools. Thus, each public school student in the sample would get added into the distribution as 1.23 students (86/70), and each private school student would get counted as .47 students (14/30).

Even though an overall sample used in norming may be very large, at any particular part of the score scale there might not be very many examinees, and thus the data distribution might be jagged even though the underlying variable is distributed more smoothly. There might be scores that no one in the norming sample obtains, although when the test is administered operationally to a larger group, those scores would be

obtained. Thus, the data collected in norming studies are often smoothed using any of a variety of statistical techniques. It is at this stage that data are interpolated or extrapolated for months from which no or insufficient data were collected.

Test Design and Development Issues

Content Coverage

Norm-referenced tests are used by school districts throughout the country, and although those districts have overlap in their curricula, they also have significant differences. Thus, content appropriate within one grade in a certain district might be most appropriate at a higher or lower grade in another district. Also, many of the decisions based on norm-referenced tests are for very high- or very low-achieving examinees, and such tests often have both material from earlier stages in the curriculum (for example, fourth-grade material on a sixth-grade test) and later in the curriculum (such as eighth-grade material on that same sixth-grade test). For both of these reasons, norm-referenced tests tend to have broader content coverage than comparable criterion-referenced tests.

Item Difficulty

A test will have the greatest accuracy of measurement for the greatest number of examinees if all test items are of middle difficulty. Middle difficulty is achieved when half of the examinees know the answer to a question. On tests where guessing can be a factor, this means that somewhat more than 50% of the examinees will answer the questions correctly. For a test consisting of five-choice multiple-choice questions, this means that 50% will answer correctly because of their knowledge and 10% (one-fifth of the remaining 50%) will answer correctly by guessing, for a total of 60% correct. Thus, if maximizing the average score accuracy was the primary concern, there would be no need for very easy or very difficult items. However, as stated before, many of the decisions based on norm-referenced tests are aimed at students at the high and low ends of the achievement continuum, and thus there is a need for accurate measurement at the ends of the score scale, resulting in easy and hard items in addition to middle difficulty items. A balance between these conflicting item difficulty requirements must be struck.

Special Topics

Issues Related to Systems of Norms for Vertically Scaled Tests

When there are multiple test forms for different grade or age levels, there are additional issues associated with developing test score norms. For example, if norms are developed separately for the third-grade form of a test and the fourth-grade form, it is desirable that regardless of the form of a test a student takes, students with the same scaled score and grade level have the same percentile rank.

Participation Rates

It is getting increasingly difficult to get a random sample of schools to agree to participate in a norming study. Many schools feel that their students are overtested. This is particularly true at the high school level. For example, in the 2005 science administration of the National Assessment of Educational Progress (NAEP), 87% of invited 4th-grade schools participated, but only 83% of 8th-grade schools and 76% of 12th-grade schools. At all three grade levels, more small schools chose not to participate. Additional schools were invited to replace those that did not participate, but it is likely that the resulting sample was unrepresentative of the nation as a whole.

Most importantly, NAEP participation rates seem to be much better than those obtained by commercial test publishers. Thus, there are many questions regarding the representativeness of published norms.

Motivation

There is concern (and some evidence) regarding the motivation of students participating in norming studies. If students (or their teachers) are not as motivated during a norming study as they are during an operational test administration, then it will appear as if the test is harder than it actually is, and norms will overestimate student relative performance.

Political Controversy

Norm-referenced and criterion-referenced tests have been caught in the crossfire of political debate on the quality of American public education. One faction feels that it is important to provide national norms to allow schools to demonstrate whether they are

performing adequately compared to schools throughout the nation. Another faction feels that norm-referenced interpretations are flawed in that if all schools improve, half of all students (by definition) will still be below average.

Percentile Data

Percentiles are ordinal-level data (for example, the difference in achievement between the 98th and 99th percentile is much larger than the difference between the 50th and 51st percentile). Thus, it may be misleading to take an average of percentiles. Instead, one should average the scaled scores associated with those percentiles and take the percentile of the average score to represent the normative performance of a group.

Individual Versus Group Norms

The average scores of groups do not vary nearly as much as the individual scores of examinees within the population. Thus, normative information to illuminate group (for example, school) performance should be based on the distribution of group averages and not on the percentile rank of the average examinee in that group. For example, a high-performing school might have an average score of 310 on some hypothetical test. A score of 310 might have a percentile rank of 75 when applied to an individual, but an average score of 310 might be better than 98% of all schools.

Adoption of Norms for Use in Customized State Tests

Standards-based tests such as those required by No Child Left Behind require a close match to state-specific curriculum frameworks. Norm-referenced tests are usually developed to be broader (and correspondingly less deep) than standards-based tests. Many states would like to provide national normative interpretations for their state standards-based test. But norms are developed for a specific test and a specific population. Some states and publishers have tried to augment norm-referenced tests with items necessary to provide the depth required of a state standards-based test, equate the state-augmented test to the nationally normed test, and use the national norms to estimate how students nationwide would have done on the state-specific test. To the extent that the curriculum frameworks differ from state to state, the results from such a process might be significantly different from those obtained if a norming study were done administering that state's test throughout the nation.

Neal Kingston

See also Criterion-Referenced Testing; Grade-Equivalent Scores; Measurement; Standardized Tests

Further Readings

Angoff, W. H. (1971). Scales, norms, and equivalent scores. In R. L. Thorndike (Ed.), *Educational measurement.* Washington, DC: American Council on Education.

Flanagan, J. C. (1951). Units, scores, and norms. In E. F. Lindquist (Ed.), *Educational measurement.* Washington, DC: American Council on Education.

Kolen, M. J. (2006). Scaling and norming. In R. L. Brennan (Ed.), *Educational measurement.* Westport, CT: Praeger.

Petersen, N. S., Kolen, M. J., & Hoover, H. D. (1988). Scaling, norming, and equating. In R. L. Linn (Ed.), *Educational measurement.* New York: American Council on Education.

O

We cannot create observers by saying *"observe,"* but by giving them the power and the means for this observation and these means are procured through education of the senses.

—Maria Montessori

OBESITY

Obesity is a medical term used to describe the amount of weight over the normal or healthy standard. It is also thought of as the amount of excess fat that one is storing on one's body. The issue of obesity was once a rarity in children and most commonly seen in adults, but in the past 20 years, the rate of obesity has more than doubled for children and tripled for adolescents. Figures ranging from 16% to 30% report that children and adolescents are either at risk or are overweight. As the incidence of childhood obesity continues to rise, so do the incidences of secondary diseases. Overweight children and adolescents are being diagnosed with "adult" illnesses more than ever, and the impact at individual, family, and societal levels is becoming greater. Some suggest that the recent and rapid increase in obesity has or will soon reach epidemic proportions.

Risk Factors

The method of determining if a child fits the criteria of obesity is the calculation of the child's body mass index (BMI). A BMI is traditionally attained by calculating a person's weight-to-height ratio. For children, BMI is calculated the same as an adult, but more information is taken into account because children are still growing, and a given BMI may be significant now but not be significant in 2 months. Instead, the BMI for children and adolescents (ages 2–19) is determined according to their sex, age (in months), and height percentiles in conjunction with a growth chart. Children under the 5th percentile of the growth chart are considered to be underweight. Children from the 5th to 84th percentile of the growth chart are considered to be in the normal range of weight. Children between the 85th and 95th percentile are at risk of being overweight. Children and adolescents over the 95th percentile are considered overweight.

For almost all children, the single risk factor of weight gain that can place a child in danger of becoming overweight is the consumption of more calories than the number of calories expended. In combination with excess caloric intake, a number of variables can compound or place some children at a higher risk of being overweight. A cause often cited by the parents and children who are obese is a medical, hormonal, or genetic condition. Conditions that can contribute to weight gain in childhood include Prader-Willi, Bardet-Biedl syndrome, Cushing syndrome, hypothyroidism, side effects of medicine, and family history. Although some children do have weight problems due to medical or genetic conditions that affect their body's ability to properly process or store calories, research has found those causes to be rare, instead

finding that the overwhelming majority of cases of childhood obesity are due to lifestyle factors.

Lifestyle factors have changed as the course of industrialization has changed and technological advances in society have been made. As technology has evolved, the roles of children have changed. Children no longer need to work on their family farms for basic sustenance or walk any distance to get necessary services. Children of industrialized nations have access to more food and have to exert little, if any, physical effort to get it—the polar opposite of reality for many children 50 years ago.

The risk factors of the modern lifestyle that contribute to childhood obesity include lack of physical exercise, poor eating habits, and home environment. Children and adolescents are less likely to engage in physical activity with neighbors or in school. Competing for their attention are video games, the television, or the Internet—all activities that require no or only a small amount of movement. More than ever, children lead a sedentary, nonactive life, with some children spending up to one third of their waking, nonschool time engaged in sedentary activities. The increased reliance on the television and computer for activities of daily life has been shown to reduce the participation in physical activity, increase the amount of snacking, and lower the metabolic rate. The sedentary lifestyle most children lead is compounded by the increased amount of snacking on non-nutritive, high-calorie food that for the past two decades has been very accessible.

Many of the snacks to which children have easy access or that they prefer are filled with sugar and fat. As children spend more time in front of the television or computer, the consumption of high-calorie foods has increased and become problematic. Advertisers have taken note that children are watching an increasing amount of television and are advertising the sugar- and fat-filled snacks directly to them. The amount of time spent at home engaging in high-risk factors has been shown to occur at a higher rate in homes with fewer fiscal and recreational resources.

Cultural and home factors include race, socioeconomic status, and family composition. Some study findings have linked factors of race, socioeconomic status, and maternal BMI with increased rates of childhood obesity. The role of race does not appear to be a factor until adolescence and seems to be the biggest issue for African American females. Although the incidence of childhood obesity appears to be the most common in Hispanic children, reports of obesity among African American children have grown the most significantly in the past 10 years. Many factors contribute to the disproportionate number of minority children who are overweight, including low socioeconomic status; limited resources (financial, educational, medical); diet; lack of awareness; and familial obesity. Another environmental factor of childhood obesity is the mother's BMI. As a mother's BMI increases, the child's BMI tends to rise. This relationship seems to have an even higher correlation when the child is in a single-parent household.

When left unchecked, lifestyle risk factors become the causes of childhood obesity. Greater awareness of these risk factors is now occurring as children develop medical problems associated with their weight. If parents are concerned that their child has one or more of the lifestyle factors that increase the risk of obesity, they should seek medical intervention for their child in order to avoid the health consequences that overweight children now face.

Consequences of Childhood Obesity

Children who are overweight are being diagnosed with diseases traditionally seen only in the adult population. These adult diseases are being diagnosed in children at an alarming rate and include high cholesterol, high blood pressure, heart disease, diabetes, arthritis, asthma, and sleep apnea. The consequences of obesity seem to be the same regardless of age. Research is indicating that more than half of overweight children have at least one risk factor for heart disease. The chronic conditions that children are dealing with can greatly affect their health as adults and their long-term mortality. As many as 80% of obese adults report being overweight as children. The long-term impact of childhood obesity is one that affects children not only physically but also mentally.

The number of children who are discriminated against or bullied by their peers because of their weight is growing. Overweight children are often ostracized by their peers for their weight, and the resulting stress can negatively affect their self-esteem, academics, and social skills. For some children, verbal abuse by their peers leads to an increase in eating as a coping strategy for the stress, which worsens the existing weight problem. The resulting decrease in self-esteem caused by the bullying and the excess overeating can place overweight children at a higher risk for anxiety and depression. The anxiety of

knowing that they might be made fun of while at school can affect their ability to focus and perform well in the classroom. The cumulative psychological effect can be and is great for some children, creating a cycle of despair. The bullying, increased eating, and emotional turmoil place an undue amount of stress on a body that is already physically stressed out. The emotional stress can have physical consequences, raising blood pressure and blood sugar levels, and causing stomach and head complaints. The emotional scars tend to be the hardest to heal, but emotional and medical support are widely available to treat the mental and physical consequences of childhood obesity.

Prevention and Treatment

As the prevalence and consequences of childhood obesity continue to grow, so do the intervention methods. Society has begun to realize that all of the conveniences and advantages that technology has provided have also provided some unforeseen negative conveniences. Obesity at any age is often described as a disease of wealth, and the easiest tool in fighting obesity is in preventing the weight gain from ever occurring.

The methods used in prevention and treatment are similar, if not the same, to varying degrees. Treatment methods include reducing caloric intake; decreasing consumption of high-fat, high-sugar foods; increasing intake of well-balanced, smaller meals; and increasing physical activity. When a child is being treated for obesity, these regulations are more stringent and structured than when a child is in prevention mode of maintaining a normal weight.

The wealth of information that acknowledges the same risk factors for childhood obesity are the same factors that, if reversed, are used to prevent obesity. The dissemination of knowledge of preventive factors of childhood obesity has been provided by the government on a federal and local level, the educational system, and medical personnel. It has been recommended that lifestyle changes be made for children who show no ill effects of a more sedentary lifestyle, before effects show up. It is vitally important for children who already possess some of the risk factors and have weight maintenance problems to change habits. The necessary lifestyle changes include increasing the amount of physical activity in which children engage and increasing the availability of nutritious and healthy food choices. The key in successful implantation of the preventive factors takes place at the family level.

Parents serve as the cook and enforcer of structure in their child's world, so it is at this level that most effective changes can be made. Teaching healthy behaviors early and consistently provides a child with the strongest protective factors from obesity. The use of lifestyle as opposed to the terminology of diet reinforces the idea of making a lifelong commitment to being a healthy person at a healthy weight. One barrier facing the implementation of prevention at this level is that many parents are not aware of the impact of childhood obesity, or how to combat it.

Parents have rated drugs, violence, smoking, and alcohol as greater risk factors to their child than obesity. Part of many parents' ignorance is due in part to their own sedentary lifestyle and lack of true understanding of how weight can affect their child. The message is being heralded by schools that have made changes in the lunches they offer and the snacks available in the vending machines. Some schools have gone as far as not allowing cupcakes to be brought in for special occasions. However, the stringency has not been as evident in schools when it comes to physical education offerings. Many schools are faced with budget cuts in which the first programs to be cut or reduced are the "extracurricular" activities, which are likely to be music education or physical education. If it isn't the budget cuts affecting the availability of physical education, it is the increase in testing that forces some schools to suspend recess if the school's scores are substandard.

The psychological and medical impacts of childhood obesity are such that the pendulum is starting to swing to acceptance of the real and growing problem. Emotional and physical intervention for those children struggling with their weight has become more responsive from many directions. Even more responsive has been the message of prevention, which has not been able to effectively reach those most at risk. With continued work at all levels, obesity may become a rarity caused by disease instead of lifestyle.

Danielle Johnson

See also Eating Disorders; Maturation; Physical Development

Further Readings

Beers, M. H. (Ed.). (2003). *Merck manual home edition* (2nd home ed., online version). Whitehouse Station, NJ: Merck Research Laboratories.

Centers for Disease Control and Prevention. (2004). The role of schools in preventing childhood obesity. *The State Education Standard*, *5*(2), 4–12.

Centers for Disease Control and Prevention. (n.d.). *Overweight and obesity: Childhood obesity.* Retrieved December 15, 2006, from http://www.cdc.gov/nccdphp/dnpa/obesity/childhood

Knight, J., & Strauss, R. S. (1999). Influence of the home environment on the development of obesity in children. *Pediatrics*, *103*(6), e85.

Mayo Clinic. (2006, March 31). *Childhood obesity.* Retrieved December 15, 2006, from http://www.mayoclinic.com/health/childhood-obesity/DS00698/DSECTION=1

Robinson, T. N. (2001). Television viewing and childhood obesity. *Pediatric Clinics of North America*, *48*(4), 1017–1025.

OBJECT PERMANENCE

Object permanence is the development of awareness that objects are separate entities that continue to exist even when one has no perceptual contact with them. Furthermore, it is the ability to retain and utilize visual images or form primitive mental images. Educational psychologists are interested in object permanence because it is seen as a developmental milestone and a significant cognitive shift. The theorist Jean Piaget introduced the concept of object permanence and its development in infants. Since his introduction of the concept around 1952, others have conducted tests to prove or disprove the notion.

Piaget proposed that infants achieve object permanence in stages during the sensorimotor period of development. This period of development has been divided into six substages. In substages 1 and 2, the first few months of life, if an infant is presented with an object and the object disappears, the infant does not search for it. In other words, the infant assumes the object no longer exists, following the rule, "out of sight, out of mind." Substage 3 (4–8 months) marks the beginning of the infant's ability to distinguish between self and others. At this stage, the infant will search for an object when it disappears if he or she were doing something with the object when it disappeared or if the object is only partially hidden from view. At this point, the infant may still think of the object in terms of his or her own actions on the object. However, when looking for the object, the infant does not usually persist in the task. As the infant starts to show clear acts of intelligence in the fourth substage, around 8–12 months, he or she will look for a fully occluded object, showing that objects now have a quality of permanence for the child. This is the point when the infant's schemes are coordinated and he or she has the skills to look for hidden objects. The infant may anticipate people and objects.

During substage 5, from 12–18 months, the child may begin to pay attention to the ways new objects or events differ from his or her present mental constructs. This is a period of discovering new means through active experimentation. This leads to the advance that children can now find an object that was hidden, exposed, and hidden again, but only if they are shown before hiding. The final substage of this developmental period occurs from around 18–24 months. At this time, a child no longer has to experiment with objects themselves and is no longer dependent on seeing or acting on the object, but instead can represent and operate on the object mentally. At this stage, the child can retrieve objects hidden in a secondary box even when he or she did not see the item moved from the original box to the other. Piaget proposed that these stages occur in order, are progressive, and are dependent on the age of the child.

Piaget conducted his experiments with infants and concluded that object permanence is achieved around eight months of age. Outward signs that an infant has achieved this developmental milestone include the exhibition of signs of separation anxiety from the primary caregiver, stranger anxiety, and delight in the game of peek-a-boo.

In more recent times, the original Piagetian theory of object permanence has been challenged. Research in the area of infant perceptual development suggests that a stable and separately distinguishable world is established earlier in infancy than proposed by Piaget. All of Piaget's tests involved infants reaching for objects, but current research suggests that reaching tests may underestimate an infant's conceptual ability, and as such, the results may instead be hampered by the immature motor ability of the infant. One popular test method is the habituation/dishabituation situation. This test only requires the infant to look at perceptual displays. The results of these studies indicate that infants as young as three and one-half months display indications of object permanence. The perceptual development research shows the possibility that infants see objects as separate from them possibly as early as birth. It has been suggested that one

explanation for an infant's failure of Piaget's tests may be due to the requirement of behaviors that the infant is not yet able to perform (i.e., pulling the cover off of an object). The exact age at which object permanence is achieved has yet to be determined through research.

Piaget's theory of sensorimotor development has been challenged by research in the field of infant perception and in the area of memory. It is believed that infants have more sophisticated perceptual and cognitive abilities at an earlier age than that suggested by Piaget's research. Some theorists suggest that there may be greater flexibility in development rather than Piaget's prescribed stages.

Lori Jackson

See also Habituation; Piaget's Theory of Cognitive Development

Further Readings

Baillargeon, R. (2004). Infants' physical world. *Current Directions in Psychological Science*, *13*, 89–94.

Mandler, J. M. (1990). A new perspective on cognitive development. *American Scientist*, *78*, 236–243.

Piaget, J. (1963). *The origins of intelligence in children.* New York: W. W. Norton.

Spelke, E. S. (1991). Physical knowledge in infancy: Reflections on Piaget's theory. In S. Carey & R. Gelman (Eds.), *The epigenesis of mind: Essays on biology and cognition.* Hillsdale, NJ: Lawrence Erlbaum.

Observational Learning

The belief that learning can occur from observing other people's behavior, and the consequences related to the behavior, is one of the most important principles of social learning theory. *Observational learning* is an important construct because a vast amount of research has revealed how social modeling can be used to decrease aggression; promote prosocial behavior such as altruism, consoling, empathy, generosity, and sharing; and facilitate the adoption of positive moral attitudes and behavior.

Prior to Albert Bandura's seminal research on observational learning, which began in the early 1960s, many researchers believed that children could not imitate behavior in the presence of an adult model, nor could they generalize imitative response

patterns transferred to novel situations in which the model was absent. Bandura and his colleagues provided counterevidence to these hypotheses and empirically showed that ideas, values, attitudes, skills, and patterns of behaviors are learned by observing models rather than caused by instincts, drives, personality traits, stimulus-response associations, or schedules of reinforcement, or governed by stage conceptions of development.

Next, an overview of the outcomes of observational learning is discussed, followed by a review of the subprocesses underlying observational learning. The focus then shifts to a description of the types of models, concluding with an overview of the characteristics of effective models.

Outcomes of Observational Learning

Social learning theorists contend that observational learning leads to five important outcomes. First, people can learn new behaviors and attitudes by watching others. Second, observing others leads to response facilitation, which motivates the observer to perform the behavior or learn when to use an already learned behavior. Third, observational learning can lead to the strengthening or weakening of inhibitions. Inhibition occurs when models are punished for performing certain actions, which then terminates or prevents the observer from engaging in imitation. Disinhibition of behaviors occurs when models perform threatening or prohibited activities without any negative consequences, thereby leading the observer to perform the same act. A fourth outcome of observational learning is that people may develop emotional reactions to situations they had never experienced in the past. A fifth potential outcome of observational learning is that people's attention is often directed to a particular object or event.

Observational Learning Subprocesses

Bandura identified four required and interrelated subprocesses of observational learning: attention, retention, production, and motivation, which must occur in order for observational learning to take place. The first subprocess, attention, specifies that people cannot learn from observing a model unless they pay attention. People select certain models based on their own needs, interests, and self-efficacy beliefs related to a given task.

The second subprocess of observational learning is referred to as retention, reflecting the importance of memory functioning in social learning theory. An individual may pay attention and observe the behaviors of a model, and yet does not perform the action until a later time. The observer must be able to remember the actions of the model and engage in the process of silent verbal coding.

Third, in order to reproduce someone else's behavior, the observer must have the necessary motor skills or physical ability to perform the behavior.

The fourth and final subprocess of observational learning is motivation, which relates to the various types of incentives to act on what the observee has learned. Motivation in observational learning is governed by the extent to which the individual values the task, and it is also influenced by principles of reinforcement. According to Bandura, the individual can receive direct reinforcement, which influences the likelihood of the person being rewarded or punished in future performances. In addition, an individual's behavior is influenced by vicarious reinforcements, or seeing a model receive some form of reward or punishment. Finally, a person may receive self-reinforcement by rewarding or punishing himself or herself based on self-evaluation of the behavior.

Types of Models and Modeling Formats

Research shows that there are three types of models that influence observational learning. First, people watch live models or other people. Observational learning, however, is not limited to watching another person. A second type are symbolic models, which may be filmed models; television cartoon characters; or characters in a novel, comic book, Internet site, or video game. A third type of model is verbal instruction or reading or hearing explicit descriptions of how to behave.

Researchers also distinguish between two types of modeling formats: mastery modeling and coping modeling. Mastery models perform flawlessly, learn quickly, and tend to verbalize statements reflecting calm, confident, and positive attitudes. In contrast, coping models may appear hesitant, show themselves making errors, and demonstrate their ability to handle challenges by relying on effort and self-reflection.

Characteristics of Effective Models

Previous research indicates that effective models typically have the following four characteristics:

1. *Are Competent.* Models need not be perfect; however, they are most often people who consistently perform at high levels of competence and have above-average capabilities.

2. *Have Prestige and Power.* Models typically are viewed as prestigious and powerful. They tend to get things done, and they are likely to effect change processes.

3. *Behave in Stereotypical Gender-Appropriate Ways.* People are more likely to observe the behaviors of others who act in ways consistent with stereotypical gender behaviors.

4. *Have Relevant Behavior.* People with prestige and power are not necessarily perceived as models. The observee must value the behavior of the potential model in order to pay attention to his or her actions.

Gypsy M. Denzine

See also Applied Behavior Analysis; Social Learning Theory

Further Readings

Bandura, A. (1977). *Social learning theory.* Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (1986). *Social foundations of thought and action.* Englewood Cliffs, NJ: Prentice Hall.

Bandura, A., Ross, D., & Ross, S. A. (1963). Imitation of film-mediated aggressive models. *Journal of Abnormal and Social Psychology*, 66, 3–11.

Bandura, A., & Walters, R. H. (1963). *Social learning and personality development.* New York: Holt, Rinehart & Winston.

Schunk, D. H. (1987). Peer models and children's behavioral change. *Review of Educational Research*, 57, 149–174.

OLDER LEARNERS

Defining older learners is a more complex process than may be assumed initially. For instance, the term *older* is misleading. *Older learner* bodes the question "Older than what?" Historically, learning has been primarily associated with the formal learning found in schools. Therefore, young learners were school aged and older learners were beyond the normal age of

schoolchildren. Reliance on chronological age alone is a fairly ineffective way to define adult learners, particularly in view of the contemporary concept of lifelong learning. The concept of lifelong learning encourages people to engage in acquiring new skills and knowledge throughout their lifetimes. This is a radical departure from the idea that a high school education was sufficient, in both knowledge and skill, to allow the graduates to assume their rightful place in the workforce and within society.

Current thought suggests that today's high school graduates may change occupations 7–10 times throughout their working lives. This need for continual learning brings an entire segment of the population, whose previous involvements were small, into the learning community. If educational psychologists create an age-based continuum of learners and arbitrarily select age 65 as the point of delineation, by 2030 it can be projected that approximately 20% of those involved in the learning process will be above age 65. They can just as easily identify those above the age of 65 as being older learners.

The use of chronological age persists as a means to identify adult learners. As in the example above, it is a simple process to arbitrarily select an age in calendar years, such as 65, and declare that those over the age of 65 are older learners. Chronological age remains simple and easy to understand, but it reveals only one component of the older learner.

Another component that may be used to define the older learner is to examine the learner in terms of his or her biological age. This is a much more complex process, and because it relates to the person's overall health, physical well-being, or the "true" age of the body, it is a construct difficult to both define and measure. However, a working estimate of a person's health may be directly related to the person's ability. For instance, some are "old for their age," whereas others are "young for their age." Although these phrases are not meant to provide empirical accuracy, they do afford insight as to a person's ability. Biological age also affects the learner's ability to sit and listen for long periods, take notes, mentally engage, and myriad other tasks associated with the learning process. These tasks may be considered a given for the younger learner, one better suited physiologically and biologically for the stress of learning, and remain a strong influencing component of the older learner.

The final age reference proposed for consideration is the concept of psychological age. Psychological age relates to the affective state, or how one feels. This feeling may relate to a positive or negative attitude toward learning specifically or in general. The important consideration is that emotionality plays a large role in an individual's learning process. People tend to become more deeply involved in things they like and less involved in pursuits they dislike. Therefore, older learners are more likely to engage in the learning process when they possess an accompanying mature psychological age, and less likely to engage when psychologically immature.

Locus of Control

As children develop, their ability to control their lives increases. Decisions initially made by the parents are shifted to the child, and as maturity increases, this locus (place) of control (autonomy) also shifts. In later mature years, the offspring may become the caregivers for the elderly parents, and thereby, roles change from dependent child to independent adult to provider for the parent. From the parents' (elderly) perspective, the locus of control has shifted from themselves to others. This new identity and all the encumbrances it brings may be a frightening transition for the older person. The loss of control can be a fearful experience.

To a lesser degree, the older learner returning to the learning environment after a long period of absence must now deal with an instructor who also represents a shift in locus of control, and this may be resented, criticized, or marginalized by the learner. A key consideration in creating a functioning learning environment is to be aware of the older student's fluctuations in the ability to control events in his or her own life. By incorporating the student into the learning process (shared locus of control) and reducing the threat of a new learning environment, the instructor can increase the likelihood of maintaining the learner's dignity, motivation, and continuation in the course or program.

Principles of Learning

From a behaviorist perspective, learning produces a resultant change in behavior. Learning may be credited with changes in attitude, values, self-efficacy, and a variety of other human attributes as well. The individual's needs, emotions, and approaches to learning shape the learning experience. Motivation and learning are inextricably linked. Motivation is not only

required to bring the adult to the learning setting, but it also influences the individual's involvement.

Cyril O. Houle promoted the idea that adult motivation, in terms of why they seek further learning, can be viewed in a three-factor relationship. Houle's Typology explains that adults frequently become involved in order to satisfy a learning objective and to seek tangible evidence of learning. This includes settings where a diploma, license, certificate, or other evidence of learning is offered at the completion of a course of instruction or a learning session. Imagine how many college students would return to their institution of higher learning after being told, "The college will no longer offer a degree, but feel free to come and learn and your reward is that you will be a better person." True, becoming a better person is implicit in the higher education experience, but without a degree at the end, it is highly unlikely that students would come at all.

Further motivation to become engaged in the learning process was identified by Houle as a learning orientation. That is, someone who enjoys the experience of learning new things. Many adult learners are quite content to move from one set of subject matter to another, gathering new skills and knowledge along the way.

The final element of Houle's Typology was explained by those who are involved in the learning process for reasons beyond the content or the learning program itself. These people enjoy the socialization of the learning experience. Getting together with a group of one's peers to engage in a learning activity can be a strong motivation for some.

Additional key elements of older learner involvement consist of needs satisfaction, positive emotionality, and for the learning experience to have an experiential component. Life experience is central to older adult learning, both in terms of what older learners bring to the classroom and in terms of what they expect to take out. This realization can bring extra care when dealing with theoretical subject matter. Pure theory is not enough to entice older learners but applied, relevant theory is more easily accepted as viable subject matter.

In numerous ways, older learners, when compared to younger learners, may appear undereducated and not prepared to engage in a contemporary learning setting. The reality is that information readily absorbed by younger learners may not have existed during the educationally formative years of older learners. These advances and the virtual knowledge explosion often

can be threatening to older learners. At times, the previously learned skills may seem insufficient or inadequate for the contemporary learning environment. For instance, the shorthand or typing skills of an older learner do not seem to fit in the age of computers. Feeling inadequately prepared for further learning may serve as a barrier for older learners.

Simple academic tasks such as reading may also have a negative impact. A reduction in their reading rate, reduced sight, and vocabulary all influence older learners' attitudes and emotions toward reading for understanding. Learners are reluctant to participate when they have a reduced self-image or a sense of being inadequate. Whether these factors are real or not, the perception of them influences their participation.

It is important to structure the learning environment so that older learners may realize early success within the learning program. Success and motivation are closely related. Motivation is not only an antecedent to learning, it is an outcome of successful learning. By realizing initial success, older learners are more likely to continue to engage in the learning process. It is worth noting that younger learners are required to learn. Whether in public, private, or homeschool settings, learning for the young is mandated by law. Adult learners are, more often than not, voluntary participants in the learning process. Even when the learning task is work related and mandated by the employer, a simple job change terminates the mandate. Therefore, the disenfranchised or dissatisfied older individual will show his or her dissatisfaction by simply walking away from the instructional setting.

Discrimination

Sometimes, the instructors of older learners unknowingly engage in discriminatory teaching practices. This is not to say that the instructors of older learners are discriminatory; rather, they are generally caring, well-meaning advocates for the older learner. Where discrimination enters is in the form of how participants are taught. The traditional delivery, set in the arena of the younger learner, typically does not directly transfer to the older learner. If the teacher of the older learner is more experienced working with the younger learner, there is a great temptation to employ similar teaching methods for both. On the surface, it may appear to be a sound practice, based on the idea that if it worked for the first group (younger), it should work for the second group (older). Unfortunately, this idea falls short in

practice. For instance, memorization seems to be a much easier task for younger learners than for older learners. The memory banks of the learner may be thought of as a file drawer or, for that matter, the entire file cabinet. Younger learners, comparatively speaking, have less stored information than their older counterparts. Experience is a critical comparative between the two. The greater the experience, the greater the number of files that exist in the file cabinet. From the limited number of files in younger learners' file system, information may be extracted with relative ease. Younger learners may have only a single drawer of knowledge files and can access a correct file quickly and easily. Older learners may have to review the entire file cabinet in order to find the required information. Ease of access to stored information, failure to link previous information with new information, and learning for the sake of repeating information (rather than applying) all have a negative impact on older learners' willingness to learn.

Speed is another factor in learning that can adversely influence older learners. When information is presented in a rapid manner, it is more difficult for older learners to keep abreast of the rate of learning. It is frequently not a issue of the subject matter being too difficult; rather, it is a matter of it being presented too rapidly.

Undereducation

The explosion of knowledge worldwide has devalued the content learned 10 or 20 years ago. Older learners may come from an era of shorthand and typing skills into a Web-based, iPod, fast-paced learning environment. The construct of basic skills should be revisited in order to ensure that the older learner possesses basic skills relevant to today's learning setting. Should a basic skill shortcoming be realized, the instructor (often younger than the learner) should anticipate the need for updating relevant skills before proceeding to more complex subject matter.

Perhaps the widest variance among older learners is found in their reading ability. Many older adults read short stories or novels, which provide an entertaining, informal source of pleasure and information. Textbook reading is often a more stressful undertaking. A reduced reading level may actually promote a reluctance to participate. Care should be taken to ensure that reading assignments are meaningful and have a direct relevance to the classroom activities.

The self-image of older learners also affects their willingness to participate. When they have a strong sense of self-worth, their experience brings meaning and they are equipped to deal with the acquisition of new knowledge and skills. Their learning outcomes are likely to be a positive reflection of effort. However, the converse is equally true. A low sense of self-worth and devalued skills and experience all contribute to a sense of defeat, often before actual learning has had a chance to be realized. It is important for the instructor to maintain an open, ongoing dialogue with the learner. Informal discussion assists the learning process and is less threatening to older learners.

The Learning Environment

If there are older learners in the classroom, there is a real possibility that starting and ending on time is a concern. Typically, older learners are products of a time when punctuality and commitment were valued attributes, more so than today. The idea of a class beginning and ending within the prescribed limits meets an expectation, which, in turn, makes the learning experience more appealing to older learners. Instructor preparedness is another critical factor for older learners. The seemingly unprepared instructor leaves unanswered questions posed by the learner, such as, "What can I learn from this person?" Again, it is a matter of expectation as older learners, more so than younger learners, have an expectation of instructor preparedness and competence. These are expectations that must be demonstrated in order for older learners to identify with and gain confidence in the learning process.

Further expectations of older learners include a need for respect and dignity. Until learners are acclimated to the learning environment, the classroom can pose a stressful, even threatening setting. Instructors should be mindful that older learners have, for the most part, achieved success in their everyday lives. Whether as workers, parents, family members, or some other endeavor, they have established and attained goals outside the classroom. Even those with previous academic success are often intimidated when they enter the classroom. Maintaining respect and dignity constitute major contributions to the learners' overall success.

The pace of learning activities has a strong influence on older learners. Sudden changes in scheduling or a rapid switch among topics within a single lesson may confuse or lose older learners. Keep individual learning sessions short, approximately 45 minutes with follow-up discussion. Reduce time pressure. Phrases

like "That is enough discussion, we have to move along" can be disconcerting for older learners who engage in discussion as a means of clarification, reinforcement, or furthering the learning process.

Regular attendance is not necessarily a strong suit among older learners. Older learners may miss a class meeting for a variety of reasons, among which are health, transportation, family-, or work-based obligations or others. An erratic attendance pattern should not be interpreted as the learner's lack of interest or ability but rather as one of the realities facing the older learner.

Timed tests may be satisfactory for younger learners but actually work to the disadvantage of older learners. Consider the knowledge of younger learners as being contained in a series of files in a file drawer. When it is necessary to retrieve a single file (recall specific information), it is a relative simple task because the files are fresh, in order, and relatively few in number. Conversely, older learners have a larger number of files (experience, previous knowledge) that may be out of sequence. In order to extract the correct response, older learners have a sufficiently difficult task in recalling from their extensive file drawer of information without compounding the difficulty with a time constraint. Elimination of timed tests, by its nature, results in higher performance levels for older learners.

Remember that formal learning (classroom learning) also has a social component. The learners will naturally gravitate to one another for support and encouragement, and may even form informal cohort groups as they progress through the learning process. These factors make the concept of group work appear more appealing than individual work. Older learners are also accustomed to working in groups, a trait acquired in both the workplace and family settings. In addition, when educators use a group- or team-learning environment, why administer individual tests? The ability to work with others is a desirable learning outcome. With that being the case, individual examination performance seems to be an evaluation method in opposition to the learning objective. Older learners are typically quick to understand this contradiction and are better assessed by means of a portfolio, research paper, group project, or some similar means of determining learning goal attainment.

In terms of classroom management for older learners, take frequent breaks and/or establish a classroom policy whereby students may enter and leave as they see fit. It may be embarrassing for some older learners to raise their hand and request permission to go use restroom facilities. A flexible entry and departure policy that demonstrates sensitivity to the learners' needs not only reduces stress and apprehension but also contributes to the overall bond of trust between learners and instructor.

Structure the class for success. Early learning success fosters later participation. Early classroom success may be as simple as a discussion activity or recognizing an individual or group for their contribution. Minimize the impact of errors in the classroom. The classroom is a desirable place to err. In the classroom, errors can be identified and corrected, furthering the prospect of success outside the classroom.

Perhaps the greatest single factor in relating to older learners is ensuring that the new knowledge to be gained is related to previously acquired knowledge and information. This linking of new information to previous knowledge and experience not only enhances the learning but also helps transform the new knowledge into practical, relevant, and functional information.

Older Learners and Younger Learners

Older learners bring a wealth of experience and understanding to the classroom. This can be intimidating to younger learners. Younger learners may also intimidate older learners, with younger learners being viewed as possessing more current knowledge and better study skills, and also being more at ease with the teaching/learning environment. With time, openness, and caring for both, older and younger learners can appreciate and capitalize on the strengths of each other. An essential element is the instructor who is cognizant of both sets of learners and willing to create a learning environment that encourages and facilitates success for all.

James E. Witte

See also Learning; Learning Strategies; Learning Style

Further Readings

Galbraith, M. W. (Ed.). (2004). *Adult learning methods.* Malabar, FL: Kreiger.

Houle, C. (1961). *The inquiring mind.* Madison: University of Wisconsin Press.

Woldkowski, R. L. (1999). *Enhancing adult motivation to learn.* San Francisco: Jossey-Bass.

OPERANT CONDITIONING

The central purpose of education is to help students to learn. The basic processes that produce learning have been investigated in the laboratory under the heading of *conditioning.* The technical term *conditioning* is used to distinguish between behavioral changes that are conditional upon—that is, dependent upon—individual experience and those that are largely independent of experience, such as salivating when a sweet substance stimulates the tongue or blinking when an object rapidly approaches the eye. This entry describes the two factors that laboratory research has identified as essential for conditioning. The entry then explores some of the implications of what is known about basic conditioning processes for learning outside the laboratory. The carefully controlled methods of the laboratory most often use nonhuman animals, but they are necessary to uncover fundamental learning processes. However, these basic processes have been confirmed in later research with humans. Following their discovery, they have been used to facilitate learning in the less controlled and more complex environments of the everyday world. Application of basic principles to behavior outside the laboratory always involves additional knowledge of the specific fields to which they are applied.

Two types of laboratory procedures have been used to study learning processes—operant conditioning and classical conditioning. The two procedures have in common the fact that each presents the learner with an environmental event (a stimulus) that already elicits some behavior. Common laboratory examples are food that elicits salivation and a puff of air to the eye that elicits blinking. A stimulus that already elicits a readily identified response provides the experimenter with a response that can be used to track the progress of learning.

The two procedures differ with regard to what kind of event precedes the eliciting stimulus. In the operant procedure, a *response* precedes the eliciting stimulus such as when a lever-pressing response is followed by the presentation of food. In the classical procedure, another *stimulus* precedes the eliciting stimulus, such as when the sound of a tone is followed by food, to use another laboratory example. In the classical procedure, the behavioral change (learning) that occurs is an increase in the strength of the elicited response to a new stimulus. Continuing with the laboratory

example, after several tone-food pairings, salivation begins to occur during the tone even before the food is presented on that pairing. In the operant procedure, the behavioral change involves not only an increase in the strength of the elicited response but also in the strength of whatever response precedes the eliciting stimulus and its elicited response. For example, after several pairings of lever pressing with food, not only does salivation increase when the animal sees the lever, but the likelihood of pressing the lever also increases.

As the foregoing examples illustrate, the classical procedure is limited to changing the stimuli that guide a response that can already be elicited: The tone comes to evoke salivation, a response that was originally elicited by food. The operant procedure is much more versatile because almost any behavior of which the learner is capable can become guided by the environment: The sight of the lever guides not only salivation but lever pressing. Using an operant procedure, the strength of any one of a large number of responses could have been changed if they had preceded the food—pulling a chain, vocalizing, running down a path, and so on. The responses that precede the eliciting stimulus are called *operants* to emphasize that they operate on the environment to produce the eliciting stimulus. Because of the greater scope of the operant procedure for changing behavior, findings from this procedure have more profound implications for understanding the learning process, particularly the conditioning of complex behavior.

Two Factors Required for Learning

Eliciting stimuli have the ability to change the way in which the environment guides behavior if two requirements are met—requirements that are technically designated as *contiguity* and *discrepancy.* Eliciting stimuli that bring about behavioral change are called *reinforcing stimuli* or simply *reinforcers.* Such stimuli reinforce, or strengthen, the ability of environmental stimuli to guide behavior. In what follows, the eliciting stimuli of concern are those that can function as reinforcers.

Contiguity

Contiguity refers to the occurrence of two events close together in time, that is, *temporal contiguity.* In the operant procedure, the effect of temporal contiguity is studied by varying the time between the occurrence of the operant and the presentation of the

reinforcer—for example, between a lever press and the presentation of food. When this time interval is varied under controlled laboratory procedures, the common finding has been that the reinforcer strengthens the environmental guidance of the response only when the reinforcer almost *immediately* follows that response. If the reinforcer is delayed for more than a few seconds after the operant, other responses have an opportunity to intervene and these responses become strengthened more than the operant (that is, the response that the experimenter has arranged to produce the reinforcer after a delay). Such findings indicate that the conditioning process is fundamentally sensitive to temporal, not causal, relations. Of course, close temporal relations are often characteristic of cause-effect relations, so the temporal contiguity requirement often provides the basis for detecting cause-effect relations between environmental and behavioral events, although the conditioning process is not infallible.

Some superstitions arise from the sensitivity of the conditioning process to temporal relations. As a laboratory example with human subjects, a college student was asked to determine what response produced points on a counter. Unknown to the student, the points were added at random time intervals independent of what the student was doing. However, because humans are always behaving, some behavior inevitably occurred immediately prior to the addition of a point. This behavior became strengthened even though it did not cause the point to be added. As a result, the response was more likely to be occurring before the next point and was strengthened further. The superstition that a particular response caused the point became established as a kind of self-fulfilling prophecy that varied for different students depending on what response happened by chance to precede the reinforcer in their pasts.

Discrepancy

In the 1970s, a second factor was identified as necessary for conditioning to occur: A reinforcer had to not only immediately follow the operant to strengthen it, but also to evoke an elicited response that was not already occurring. In technical terms, to function as a reinforcer, an eliciting stimulus must produce a behavioral change (or *behavioral discrepancy*) as well as occur close in time to the operant. As a laboratory example, if pressing a lever is first followed by

food when a tone occurs, and later in training, a light is added to accompany the tone, then the light never acquires the ability to guide lever pressing when the light is presented by itself. Because of prior conditioning to the tone (which caused the tone to guide salivation as well as bar pressing), conditioning to the light was blocked because the animal was already salivating when the light was introduced. That is, there was no discrepancy between the behavior occurring during the light and the behavior elicited by the food.

Stated in everyday language, laboratory studies have discovered that learning occurs only when behavior is immediately followed by an event that "surprises" the learner.

Implications for Education

Some of the many implications of operant conditioning for education are sketched here, which research supports.

Initial Learning (Acquisition)

Except for the very simplest behavior, the basic procedures used in the laboratory to study operant conditioning must be elaborated to condition more complex behavior. Most real-world behavior consists of a coordinated sequence of several distinct responses, as when writing even a single word. In such cases, at the outset of learning, the presentation of a reinforcer—such as praise from a teacher—cannot be delayed until the completion of the entire sequence of responses. The writing of each individual letter or part of a letter must be reinforced first, with subsequent reinforcers gradually withheld as successive responses approximate ever more closely the entire sequence. This technique is called reinforcement by *successive approximations* to the final behavior.

Conditioned Reinforcement

Most reinforcers for complex human behavior are not innate elicitors, such as food evoking salivation. Praise from a teacher is an example of a learned or *conditioned reinforcer.* Conditioned reinforcers are stimuli that have acquired their ability to evoke behavior (e.g., praise evokes smiling) through being paired with other stimuli that are already elicitors. Research has determined that the principles that

describe the action of conditioned reinforcers are essentially the same as those for innate elicitors.

Maintenance of Behavior (Intermittent Reinforcement)

Initial learning is facilitated when every occurrence of the target behavior is immediately followed by a reinforcer. However, behavior is better maintained when the frequency of reinforcers is gradually reduced during training. For example, a child who is learning to write might receive praise at first for every well-formed letter, but later, praise should be withheld until an increasingly large number of letters has been produced. When reinforcers do not occur after every instance of the target behavior, the procedure is called *intermittent reinforcement*. Behavior is well-maintained by intermittent reinforcement and persists for a longer time if, for any reason, reinforcers are later missing.

Reducing the frequency of reinforcers for the target response also keeps the reinforcer "surprising." If the response evoked by the eliciting stimulus is fully "expected," it will no longer function as a reinforcer. Thus, when praise is dispensed frequently, especially when performance is marginal, praise loses its effectiveness as a reinforcer.

Environmental Guidance of Behavior (Generalization and Discrimination)

Reinforcers do not merely strengthen behavior; they strengthen behavior in the presence of the particular environment in which the operant produced the reinforcer. The reinforced behavior will later occur during other environments only to the extent that other environments share features in common with the training environment. This is known as *stimulus generalization*. For learning to generalize from the classroom to the wider world, the classroom should contain features that overlap with those of the environments in which the behavior is intended to occur. Sometimes, however, generalization of learned behavior must be restricted to particular environments, even when those other environments share features in common with the training environment. For example, calling out a friend's name on the playground may be reinforced by the friend's attention but not in the library when the friend is studying. The restriction of reinforced behavior to a particular environment is called *stimulus discrimination*. Stimulus discriminations are acquired when behavior is reinforced in the presence of one environment but not in the presence of other environments, even if the other environments contain some common features.

Applications of Operant Conditioning in Education

A number of successful efforts have been made to explicitly apply principles and findings from the study of operant conditioning to educational practice. Among these are the *Personalized System of Instruction* (PSI), *Precision Teaching* (PI), and *Direct Instruction* (DI). Although the methods differ somewhat, they share a number of features in common. Among these are immediate feedback (reinforcement) for the performance of individual students, frequent objective measures of performance, specific procedures to implement the method, and reliance on instructional research. For example, in PSI, each student moves at his or her own pace with progress to the next step not occurring until the student has mastered knowledge at the present step. Thus, individual differences between students in grades are replaced by individual differences in rates of achievement of mastery. Objective assessments indicate that these methods foster student achievement that is superior to standard teaching methods for students of a wide range of abilities. As one example, the superior achievement produces, as a by-product, greater student self-esteem, even greater than alternative procedures explicitly designed to foster self-esteem. Nevertheless, operant conditioning teaching methods remain controversial because they are seen as at odds with mainstream educational philosophy.

John W. Donahoe

See also Applied Behavior Analysis; Classical Conditioning; Learning; Reinforcement

Further Readings

Binder, C., & Watkins, C. L. (1990). Precision teaching and direct instruction: Measurably superior instructional technology in schools. *Performance Improvement Quarterly, 3*, 74–96. Retrieved from http://www.behavior.org/education

Chance, P. (2003). *Learning and behavior.* Belmont, CA: Thompson/Wadsworth.

Donahoe, J. W., & Palmer, D. C. (2004). *Learning and complex behavior.* Richmond, MA: Ledgetop. Retrieved from http://www.lcb-online.org

Skinner, B. F. (1953). *Science and human behavior.* New York: Macmillan.

Watkins, C. L. (1997). *Project follow through: A case study of contingencies influencing instructional practices of the educational establishment.* Cambridge, MA: Cambridge Center for Behavioral Studies.

P

When you teach your son, you teach your son's son.

—The Talmud

PARENTAL EXPECTATIONS

Psychologists and sociologists have long been interested in parents' expectations about their children's achievement in school and their future occupational attainment. Although the term *expectations* has been defined in various ways in the literature, most current research tends to characterize parent expectations as realistic beliefs or judgments that parents have about their children's future achievement. Parent expectations can be distinguished from aspirations, which refer to idealistic hopes or goals that parents may form regarding future attainment. In some studies, parents are asked directly about their expectations, but it is also common to ask students to articulate what they perceive to be their parents' expectations.

Early work on parental expectations was conducted by sociologists interested in understanding how a family's socioeconomic status (SES) is transmitted from one generation to the next. These researchers sought to pinpoint the transmission mechanism involved and began including parental expectations to their models as a mediator between family SES and adolescents' educational and occupational attainment. Significant family members were viewed as "definers"—those who held expectations and conveyed them directly—or as "models"—those who influenced students indirectly through their own aspirations or level of attainment.

Subsequent theorists seeking to elaborate a process model by which families affected student achievement began to examine how parents expressed and communicated academic expectations to their children. Four major questions have been posed in contemporary research on parental expectations: Do parental expectations influence children's achievement in school? What are the mechanisms linking parental expectations and student achievement? What factors determine the nature of parental expectations? What factors determine the strength of the relation between parental expectations and student achievement?

Relation of Parental Expectations to Student Achievement

The relation between parental expectations and student achievement has been established in many empirical studies. Parents who hold high expectations for their children's grades or performance on examinations tend to have children who subsequently receive higher grades, achieve higher scores on standardized tests, hold higher aspirations for future educational and occupational attainment, and persist in their schooling. Several meta-analyses have found that parental expectations are the strongest family-level predictor of student achievement outcomes, exceeding the variance accounted for by other parental beliefs and behaviors by a substantial margin. By controlling on the effects of early

achievement, the results of longitudinal studies offer particularly powerful proof that parental expectations are a causal determinant of student outcomes.

Causal Mechanisms Linking Parental Expectations to Student Achievement

Early accounts by sociologists postulated that parental expectations signaled information to students about their capabilities that was then incorporated into students' aspirations for achievement. Social cognitive theorists have similarly argued that high parental expectations contribute to scholastic achievement by raising students' academic self-efficacy and aspirations, as well as by bolstering their confidence in counteracting peer pressure to engage in detrimental pursuits.

A second line of work, emerging in the 1960s and 1970s, introduced the notion of "achievement press," arguing that parental expectations create a psychological demand to realize those expectations. The notion that parental expectations create a pressure to achieve is reflected in current work examining whether high expectations may sometimes cause emotional stress in children. A third line of inquiry has focused on how parental expectations catalyze parent efforts to engage in specific forms of teaching, monitoring, and supervision of their children's learning; to shape the child's academic environment; or to form connections with the school. A fourth perspective emphasizes the connection between high expectations and parents' emotional support for student achievement. Following James Coleman's notion of social capital, these theorists argue that positive parent–child interactions express themselves in the form of high parental expectations, thus establishing parental expectations as the mediator rather than an exogenous determinant of nurturing relations.

Determinants of Parental Expectations

At the level of the individual family, the most salient factor affecting the nature of parental expectations is prior student performance. Students who have been performing at a certain level are, to some extent, expected to continue at that level. However, subjective factors also come into play as parents formulate expectations for future performance. Just as high-achieving students may lack a sense of their own academic capability, so parents may miss the relevant cues and make unduly optimistic or pessimistic predictions concerning their children's future.

Factors operating at the level of racial or ethnic group affect parental expectations as well. Many studies have documented high academic expectations held by Asian American parents, relative to other ethnic groups. Most researchers attribute this pattern of high expectations to cultural beliefs about the importance of education. Others argue that education is seen as the only viable means for Asian Americans to secure economic security in the face of discrimination. Parental expectations also vary by parent SES, with mother's education level being particularly influential. The fact that SES predicts parental expectations after controlling for student prior achievement suggests that these SES differences are not simply a reflection of class-based achievement patterns.

Contribution to Salience of Parental Expectations

Many investigators have sought to understand the conditions under which parental expectations are particularly salient. Four factors contribute to the strength of the relation between parental expectations and student achievement. One factor is the closeness of the parent–child relationship. Students who interact with their parents frequently and who have a sense of mutual trust and commitment may be more affected by their parents' expectations. A second factor is the age of the child. Parental expectations are less strongly associated with student performance in elementary than in secondary school, perhaps because parents have obtained more information about their children over time. A third factor pertains to the completeness and accuracy of information that parents receive about their children's achievement. Parents may form erroneous expectations if the information is difficult to obtain or is not provided in a clear, understandable format. A fourth factor refers to parental resources. Parents with resources to act on their expectations with supportive actions may effect a change in their children's achievement that is in line with their expectations, but in families that lack such resources, children's performance may decline in spite of high parental expectations because parents are unable to provide needed support.

Susan D. Holloway and Yoko Yamamoto

See also Parenting; Parenting Styles; Social Development

Further Readings

Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (1996). Multifaceted impact of self-efficacy beliefs on academic functioning. *Child Development*, *67*, 1206–1222.

Haller, A. O., & Portes, A. (1973). Status attainment processes. *Sociology of Education*, *46*, 51–91.

Hao, L., & Bonstead-Burns, M. (1998). Parent-child differences in educational expectations and the academic achievement of immigrant and native students. *Sociology of Education*, *71*, 175–198.

Juang, L. P., & Silbereisen, R. K. (2002). The relationship between adolescent academic capability beliefs, parenting and school grades. *Journal of Adolescence*, *25*, 3–18.

Okagaki, L., & Frensch, P. A. (1998). Parenting and children's school achievement: A multiethnic perspective. *American Educational Research Journal*, *35*, 123–144.

Sandefur, G. D., Meier, A. M., & Campbell, M. E. (2006). Family resources, social capital, and college attendance. *Social Science Research*, *35*, 525–553.

Parenting

Parenting has become a commonly used term. As both a noun and verb, it is now applied to the many aspects of caring for a child from birth to adulthood; however, the word did not emerge in public usage in America until the 1950s. It was preceded by the term *parentcraft*, which was used in Britain in the 1930s to refer to learning the skills related to a wide spectrum of practical family matters. Even today, with increased emphasis on the importance of parenting and the prevalence of available child-rearing information, there is no general agreement on what *parenting* specifically means.

Rather than a unitary concept, the term invokes a host of assumptions and expectations and encompasses a cluster of meanings. Broadly speaking, parenting might best be defined as the process of rearing children; providing for their health, education, and well-being; and ensuring their cognitive, social, moral, spiritual, and emotional development. It involves the holistic preparation of children to become successful, independent adults.

Background

Before the 20th century, information on how one should parent was generally found in the counsel of older, experienced mothers and religious leaders. It was not until the work of pediatrician L. Emmett Holt, who wrote the *Care and Feeding of Children* in 1894, that a scientific approach to the study of parenting began to take hold. Holt's practical, stern book that recommended strict recipe-like control and routine remained popular for years, with new editions appearing until 1934, by which time numerous child psychologists, other doctors, and child study experts began producing a variety of works on parenting.

The focus of these early parenting books was a mix of psychology and medicine. John Watson, who established the behaviorist school of psychological thinking, brought public attention to ways parents mold a child's development. His book, the *Psychological Care of Infant and Child*, published in 1928, was based on his ethological studies of animal behavior, from which he drew extravagant conclusions that children could be molded into any kind of person a parent desired through consistent behavioral techniques. He admonished parents to treat children with emotional detachment but also with respect. This approach opened the door for questioning the puritanical, autocratic belief that corporal punishment was a necessary component of strict discipline; however, at the same time, he criticized parental responsiveness, warning that warmth and affection would "spoil" a child. Watson's belief that parenting is a science and should be regarded as an ongoing experiment in successfully rearing children became widespread through popular magazines, radio broadcasts, and speeches to educators and physicians. The U.S. Department of Health printed his views in parenting pamphlets.

Pediatrician Benjamin Spock opposed Watson's approach to scheduled, autocratic parenting. His highly influential book, *Baby and Child Care*, published in 1946, advised parents to regard each child as an individual and to provide loving guidance within an environment of firm, consistent limits. As a result of his writings, child-rearing was viewed flexibly, within the context of changeable family dynamics.

Throughout the last half of the 20th century, there has been a proliferation of eclectic, often contradictory parenting advice directed to those searching for ways to optimally rear a child. Theory and research has become segmented, with separate foci on specific areas such as discipline, medical issues, developmental problems, and developmental stages. This trend has provided deeper insight for professionals in the field, but it can be confusing for parents who rarely have the depth of knowledge or experience required

to synthesize across studies. The highly publicized separate reports of factors supporting children's optimal growth may contribute to an increase in parental anxiety, particularly in those who become parents later in life and those who have fewer children.

Traditional and Evolving Parenting Roles

Because the word *parenting* is not representative of a static conception but refers to different behaviors and practices, it has developed distinct meanings within different contexts and with different groups. It is important to acknowledge that parenting can be seen from a variety of perspectives representing a variety of dimensions, ages, and stages. Parenting behaviors, taking into account all aspects of a child's social, emotional, physical, moral, spiritual, and educational development, are influenced not only by the culture in which the parent belongs but also by the individual family structure and the community in which the family lives.

In Western societies, parenting obligations have traditionally depended on biological relatedness. Now, from a legal perspective, parenting has become a contentious issue as a result of the changing nature of family relationships. The increase in surrogacy, foster care, and nontraditional adoption has presented new issues requiring resolution. Questions of the rights and responsibilities of a biological parent (producing a child) versus those of a functional or social parent (rearing a child) make it essential that the areas of custody, care and support, and personal identity be clearly identified. When used in legal cases such as adoption, custody, the assignment of rights and responsibilities, or making educational decisions, the terminology referring to parenting is subject to state or local judicial interpretation. Clearly drawn legal parenting job descriptions and agreements can be used to delineate roles and responsibilities for blended families, encourage effective co-parenting, and minimize conflict among divorced and divorcing spouses.

In 1989, the United Nations met to develop a set of minimum standards to protect the rights of children. Members of the Convention on the Rights of the Child universally agreed on an absolute right to a minimum standard of care and support for those under the age of 18. Parental responsibilities, as adopted in Article 27(2), state, "The parent(s) or others responsible for the child have the primary responsibility to secure, within their abilities and financial capabilities, the conditions of living necessary for the child's development."

The Supreme Court of the United States has consistently upheld the tenet that parents have the right to direct the basic upbringing and education of their children. Important precedents have been established that address religious training, medical decisions, and educational opportunities and requirements. Since the *Meyer v. Nebraska* case in 1942, the right to rear one's children has been deemed "essential," and the integrity of the family has been considered protected by the due process clause of the Fourteenth Amendment. In 2000, the Court wrote a landmark decision (*Troxel v. Granville*) affirming the "fundamental right of parents to rear their children." Subsequent decisions have held that these "fundamental rights" include the custody, care, and control of their children. Interpretations of these opinions are complicated by the changing demographics of parents.

The Changing Demographics of Parents

A review of U.S. Census data reveals significant ongoing changes in the makeup and structure of today's families. Many of these changes have major implications for parenting, particularly in relationship to the educational of the nation's children. Each census shows the increasing cultural diversity of the United States, bringing with it significant variations in family structure, child-rearing practices, and socialization patterns.

Single-parent households now constitute nearly one third of homes with children under age 18. More than 40% of children in single-parent households live in poverty compared with less than 10% in two-parent homes. There are more than 2.5 million grandparents raising children. Moreover, the extreme youth of many new parents continues to be striking, with father-absent teenage births remaining much higher than those of just a few generations past. Even two-parent families are much likelier to have both parents working than ever before, limiting time available to spend with children. This poses significant pressures related to child-care issues, including additional responsibilities and duties related to schools.

According to the Stepfamily Association of America, more of half of current Americans are, have been, or will be in one or more stepfamily situations. In addition, as noted by the Children's Defense Fund, more than 4% of the nation's children live with neither

biological parent. There are many children in adoptive, foster care, or grandparenting environments, and as the number of same-sex households continues to grow, many children are being reared by gay and lesbian parents. Effective parenting with regard to the development, adjustment, and psychological well-being of children must take these different family configurations into account.

In addition to changing family structures, the mobile lifestyle of Americans presents other challenges for parenting. Nearly 15% of Americans move annually. This puts a burden on both families and schools trying to provide stability and educational continuity. This mobility has created long-distance families, which tends to diminish the sharing of child-rearing wisdom across generations and within extended families.

The cultural diversity of the United States presents different parenting concerns. The 2000 U.S. Census data show the Latino population is growing four times faster than the general U.S. population. People of Hispanic heritage account for half of all U.S. births, and one in five U.S. teens identify as Hispanic. Hispanics comprise 13.5% of the population. Comparatively, the African American population counted a slower gain (3.1%), and the Asian population approached a 10% growth rate. The cultural history of each of these individual groups, as well as many other immigrant groups and native peoples, brings richly varied expectations and traditions to the tasks of parenting. These population shifts warrant the review of generalizations and assumptions of parenting roles based on previously drawn heterogeneous models.

Parenting Styles and Attitudes

Educators and psychologists have become increasingly interested in the ways parents can influence a child's development. This has led to substantial research on parenting styles and practices. Studies focus on parental values, expectations, personality and temperament, as well as broad patterns of parenting style.

There are four widely acknowledged parenting styles, which were first identified in the work of Diana Baumrind in the early 1990s. These styles have been linked to predictions of child well-being and achievement. The authoritative style is the most common approach of parents in the United States, regardless of cultural background.

Authoritative parenting is child-centered. Authoritative parents are highly involved and actively interested in their child's life. They tend to value open communication and choice and support their child's growing independence. At the same time, authoritative parents do not shy away from clear expectations or appropriate, not punitive, discipline. Current literature may include the term *democratic parenting* when referring to this general style. Children reared by parents with this parenting style are encouraged to make choices, see the consequences of their choices, and creatively cope with change, which makes authoritative parenting a good match with the demands of a quickly evolving society. On-going research points to the success of authoritative parenting in raising secure, responsible children.

In contrast to authoritative parenting, *authoritarian parenting* is adult-centered. Authoritarian parents exercise strict control over their child and tend not to value open communication. They expect obedience and may use physical punishment as a preferred method of discipline. This type of parenting demands adherence to established norms, provides instruction through a master–pupil or adult–child model, and is opposed to change. Children reared by parents using an authoritarian parenting style may achieve well academically, but they may also become rebellious or fall victim to peer pressure.

Both authoritative and authoritarian parents have high expectations for their children. One primary difference is that the authoritative parent uses explanation and discussion, whereas the authoritarian parent uses strong, unquestioning control.

The third type of parenting style is called *permissive*; these parents are predominantly lenient and indulgent. Like authoritative parents, they allow choice and independence and do not provide clear boundaries, guidelines, or expectations. These parents value nonconformity. Whereas children raised by permissive, nondirective parents may become inventive and creative, they often have problems when working in structured environments such as classrooms or offices, because they do not have experience with appropriate behavioral expectations. In addition, children reared in indulgent homes score higher on measures of self-esteem and social skills but are more likely to have difficulty adjusting to school.

The fourth parenting style, *neglectful parenting*, is represented by uninvolved parents who demand little or nothing from their child, are not responsive or aware of their child's needs, and do not actively participate in their child's life.

There have been years of research into parenting styles; however, it remains a vital area of study especially in light of changing family demographics.

Studies of Parenting

Topics frequently included in studies about parenting cover the direct ways parents impact the development of their children. Research into parenting practices can be divided into the relationship between positive and negative influences on child outcomes. Positive practices include supportive involvement and consistent, caring monitoring.

The bulk of studies look at the ways into which negative parenting practices, such as corporeal punishment, lack of support, or poor supervision, put children at risk for cognitive, behavioral, social, and health problems. On the other hand, there is research into parental fostering of talent development, resilience, and achievement. Positive parenting has been shown to foster adaptive behaviors as well as direct and indirect effects on adjustment.

Measures used to evaluate aspects of parenting are predominantly self-report questionnaires and observation scales. These rate monitoring and discipline techniques, relationship and interaction quality between parent and child, family context (including adaptability, cohesion, organization, and stress), and involvement level of parent at home and at school.

Research into factors leading to learning and school achievement has shown that there is a positive link between the level and quality of parental involvement and benefits for children. Although this research offers compelling evidence to support ongoing studies, there are some inherent limitations. Much of the work into issues related to the relationship between parenting and education has been carried out using experimental designs such as ethnographies, case studies, or other qualitative studies that do not allow cause and effect conclusions.

Parenting Behaviors

Jeanne Brooks-Gunn and Lisa Markman have proposed seven aspects of parenting: nurturance, language use, discipline, materials in the home, monitoring, management of the home, and, direct teaching of skills to children. Several of these dimensions, such as language, materials in the home, and the teaching of skills, have been directly linked to school achievement. Other parenting practices, such as discipline, nurturance, and monitoring, are linked to potential behavior problems. Programs for parent training have been shown to have an effect on parenting behaviors, as they help parents build awareness and interaction skills.

Parental Involvement in Education

The right of every child to a uniform, consistent education is a hallmark of America. The expansion of public education has been praised as the most successful progressive movement of the 19th century. Until the 1840s, most elementary schools in the United States were cooperatives, organized and run by groups of parents. Children were also taught to read in Sunday schools using the Bible. The "common school" movement took hold as community leaders began to see the practical and social benefits of a basic education for all children. By 1870, each state in the United States provided free elementary schooling, marking a profound change in the relationship and perceived responsibility of parents for the education of their children.

Parents and families, who had been the center of their children's education, became increasingly peripheral to the conception of education in the newly emerging public schools and school practices. In addition, the reliance on schools as centers of education tended to blur the distinction between a child's education and formal schooling. Parents were encouraged to leave the education of their child to professional educators. However, it is understood today that education, as a process of lifelong learning, takes place in all areas of life, not just in school. Therefore, home experiences, including parental input, provide important influences on learning, both positive and negative. For example, when children play, listen to their parents and friends, read books, watch TV, and engage in other activities, they are becoming educated in a variety of ways.

Concerned parents noted the dangers of separating school and home. The formation of the Parent Teacher Association (PTA) was an attempt by parents to encourage collaboration and communication among all stakeholders. In 1897, mothers and fathers, as well as teachers, workers, and politicians, gathered for the very first PTA meeting. By the early 1900s, the PTA had grown dramatically in numbers and in strength, becoming a fixture in American life and schools. The association's efforts led to the creation of kindergarten

classes, a juvenile justice system, and child labor laws. Other achievements of the PTA included the founding of the hot lunch program, mandatory immunizations for schoolchildren, physical education classes in public schools, and sex education for preteens, as well as increased wages for teachers.

The history of the PTA, an example of how parents usefully intervene in the education of their children, dramatizes one of the underlying tensions in American education. Overall, school administrators initially were receptive to the PTA and the involvement of parents. However, as PTA members tried to become involved not only in the social welfare of the students but also in the educational content of programs, tensions developed between the predominantly male administrators and school board members, and the predominantly female membership of the PTA. Beyond considering such issues as whether hot lunches should be available to their children, women desired to influence issues such as when reading should begin and what textbooks should be used. Some administrators resisted this extension of the PTA's work.

As the PTA grew and became a formidable institution of its own, other forms of independent parent teacher organizations (PTOs) began to emerge. PTOs are groups that choose to remain independent of a national organization, choosing to respond to local issues and needs. These local groups point to one of the critical trends describing parent involvement in the schools: Over the past several generations, the PTA has consistently lost members. From a record high of 12.1 million members in 1962, membership dropped to about 6 million, despite record-high school enrollments.

Explanations for this decline include factors that contribute to limited involvement of parents in schools: economic pressures on families that may require both parents to work, the growth of single-parent households, the increased mobility of Americans, and the resulting geographic distance between extended family supports.

National Recommendations

Current federal legislation recognizes the importance of parental involvement. The No Child Left Behind (NCLB) Act of 2001 contains provisions that require schools to provide information about student progress in understandable and accessible language for parents. The law also requires that schools offer professional development activities to train educators about ways to involve, and work with, parents in support of each child's education.

In the landmark 1983 report *A Nation at Risk*, after making numerous recommendations for educational reform, the members of the National Commission on Excellence in Education directly refer to the crucial role of parenting in education:

> *To Parents:* You know that you cannot confidently launch your children into today's world unless they are of strong character and well-educated in the use of language, science, and mathematics. They must possess a deep respect for intelligence, achievement, and learning, and the skills needed to use them; for setting goals; and for disciplined work. That respect must be accompanied by an intolerance for the shoddy and second-rate masquerading as "good enough." You have the right to demand for your children the best our schools and colleges can provide. Your vigilance and your refusal to be satisfied with less than the best are the imperative first step. But your right to a proper education for your children carries a double responsibility. As surely as you are your child's first and most influential teacher, your child's ideas about education and its significance begin with you. You must be a *living* example of what you expect your children to honor and to emulate. Moreover, you bear a responsibility to participate actively in your child's education. You should encourage more diligent study and discourage satisfaction with mediocrity and the attitude that says "let it slide"; monitor your child's study; encourage good study habits; encourage your child to take more demanding rather than less demanding courses; nurture your child's curiosity, creativity, and confidence; and be an active participant in the work of the schools. Above all, exhibit a commitment to continued learning in your own life. Finally, help your children understand that excellence in education cannot be achieved without intellectual and moral integrity coupled with hard work and commitment. Children will look to their parents and teachers as models of such virtues.

Parenting Programs and Emerging Models

Many educators are not convinced that parent involvement can make a difference, nor do they understand how parent involvement can be a significant help.

A majority of teachers receive little preservice education about how to work with parents and involve them productively in the lives of their children. Because the rates of return on educational experiences are greatest for young children, it makes sense for teachers and administrators to reach out to inexperienced parents of children in the early grades. This is where many of the most productive efforts involving parents in schools have been focused in the recent past.

It should also be noted, however, that a major impetus for the development of alternative and charter schools has been their willingness to give parents a greater role in the educational process than many feel is unavailable in the public schools. For example, the Central Park Elementary School (CPE) is a highly successful alternative school that was founded in the 1970s by Deborah Meier to emphasize active learning. At CPE and the schools that grew out of it, Meier succeeded by fostering a democratic community that gave teachers greater autonomy in the running of a school along with giving parents a strong voice. This promoted a family-oriented system. The CPE Secondary School, with its parent–teacher teamwork, has been praised as a model of urban education reform, characterized by an environment of nurturing adults with high standards. Similar chronicles of successful schools, both public and private, emphasize the central role that parents play in the culture of such schools and the lives of their students.

Parenting in the 21st Century

Parenting has never been more challenging than it is today, with the entire range of issues surrounding the raising of a child intensified by an increasing variety of family configurations, cultural demographics, and educational choices. Many parents feel that much more is demanded of them than was demanded of parents in previous generations to successfully raise their children in such a rapidly changing world. Yet along with all the increased demands, there are greater resources available to assist parents than ever before.

Legal precedents will continue to define parental roles and allowable parental behaviors in the years ahead, but one of the guiding principles of parenthood might be that there is no rule for applying a rule. Establishing and developing networks of support between families and communities is as necessary as cognitive understanding of the tasks of parenting and

the resources that may be available through contemporary research or new information technologies.

Robin Schader

See also Ethnicity and Race; Family Influences; Individual Differences; Parenting Styles

Further Readings

Baumrind, D. (1971). Current patterns of parental authority. *Developmental Psychology Monographs*, *4*, 99–103.

Belsky, K. (1984). The determinants of parenting. *Child Development*, 55, 83–96.

National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform. A report to the nation and the Secretary of Education, United States Department of Education.* Retrieved from http://www.ed.gov/pubs/NatAtRisk/recomm.html

Spock, B. (2004). *Dr. Spock's baby and child care.* New York: Pocket Books.

Parenting Styles

For several decades, developmental psychologists have explored the issue of how to characterize "good" or adaptive parenting, hoping to identify those parenting behaviors that promote healthy child development. Developmental researchers have investigated distinct parenting styles, or general patterns of caregiving, that have been linked either positively or negatively with children's overall functioning. These different styles of parenting were first explored in laboratory and naturalistic observations of parent–child interactions during the 1960s. Since that time, clinical researchers have homed in on specific parenting practices reflective of these broad parenting styles in an effort to learn how and why maladaptive parenting might lead to particular childhood disorders. In recent years, research into effective parenting has focused largely on two global constructs that appear to positively influence child development: *responsiveness* and *demandingness.*

Responsiveness

Responsiveness typically refers to the interactive synchrony between a caregiver and child. Responsive

parents are able to read and respond to their children's needs with sensitivity and warmth. They are emotionally engaged with their child and consistently meet their child's needs. Responsive parents also value reciprocity and consider the child's wishes and needs. In this way, responsive parents make room for genuine exchange with their child. Being a responsive parent requires the ability to think flexibly, to problem-solve, and to adjust expectations in light of children's evolving needs and emerging personalities.

Although parental responsiveness is important across the span of childhood and adolescence, much of the research into parental responsiveness has focused on the relationships between mothers and their infants and toddlers.

Over time, sensitive care contributes to a child's development of secure attachment. For this reason, responsiveness toward children is essential, even in interactions with very young infants. In research into parenting of infants, responsiveness is operationalized by measuring features such as a mother's affection and skill when handling her baby; her sensitivity to the baby's crying and nonverbal signals; and her accessibility to, and consistency with, the baby. Additional measures of responsiveness investigate the degree to which a mother accepts the individuality of her child and how adaptable she appears in interactions with her child. As children mature and their needs change, parents need to learn to respond in new ways. Thus, researchers who study the relationship between parents and older children will operationalize responsive behaviors differently. Evidence suggests that mothers who are consistently responsive tend to experience greater harmony and cooperation in their interactions with young children, and these children tend to score high on measures of warmth and self-regulation.

Parents who are unable or unwilling to tune in to their child's needs may exhibit disengaged, inconsistent, neglectful, or even abusive parenting behaviors that are detrimental to the child's emotional, social, moral, and intellectual development. Jude Cassidy and colleagues have shown that children of rejecting, unavailable, or unresponsive parents are at increased risk for developing insecure attachment. For example, depressed mothers tend to score low on measures of responsiveness, heightening their children's risk of developing insecure attachment. Cassidy's research has demonstrated that when depressed mothers are separated briefly from their children, mother and child

tend to avoid eye contact at reunion, and the child may even react with sadness and avoid the mother. Similar findings have linked maternal inconsistency to mother–child interactions characterized by low warmth and awkwardness.

On the opposite end of the spectrum, parents who are overly responsive may be overprotective and intrusive, which can strain their relationship with their child as well. Empathy and care are important elements of parent–child bonding, but part of being a responsive parent involves learning when to let go. Warm, engaged parents support children's exploration of their environment, which in turn provides the children with learning opportunities and the chance to develop their independence. Intrusive or overprotective parents, on the other hand, may restrict young children's ability to explore and learn independently.

Demandingness

Demandingness is another factor that can foster positive development in children. *Demandingness* describes how parents monitor and discipline their children. Researchers who investigate parental demandingness might examine the closeness with which children's behavior is monitored, the types of limits that are in place within a family, the dialogue surrounding rules and values, the consequences of rule violation, or the consistency with which rules are enforced. Like research into parental responsiveness, studies examining parental demandingness are shaped by the age of the participants' children, as parents demand very different behaviors of adolescents, for example, than they do of toddlers. Appropriate levels of demandingness have been associated with positive developmental outcomes for children.

From the time children are very small, parents demand certain behaviors of their children. Parents typically begin setting limits and rules with children by their second birthday. As children grow, parents expect more cooperation and better adherence to rules. Age-appropriate rule setting is essential to the child's development of self-regulation and moral reasoning skills. The consistent enforcement of appropriate rules is part of an ongoing learning process that teaches children how to self-regulate and independently form moral judgments.

In the absence of demandingness, rules may not be set and enforced appropriately. Children whose parents demand little of them may show poor self-control

and moral reasoning skills. However, when parents are excessively demanding, the same results may occur. Paradoxically, parents who demand total obedience may well encounter high rates of problem behavior in their children. Overly demanding parents may also be highly confrontational, which can be overstimulating for children and may provoke further noncompliance. Grazyna Kochanska and colleagues found that power-assertive parents often rely on tactics such as coercion, pressure, harsh insistence, and criticism as means to extort their children's compliance. These researchers have linked such parenting techniques to negative outcomes for children, including poor self-regulation and increased aggression and conduct disturbances.

A final concern that emerges when parents are overly demanding pertains to the content of their demands. Parents need to be reasonable in setting age-appropriate standards for their children. Overly demanding parents may have excessively high expectations for their child. When disappointed by their child's inability to meet these expectations, parents may overreact in harsh or punitive ways. Michael Lorber and his colleagues have found that these overactive responses, in turn, may contribute to the early onset of childhood externalizing problems.

Parenting Styles and Child Outcomes

Based on an extensive body of literature, psychologists have reached the conclusion that "good" or skillful parenting is typically characterized by a balance of reasonably high levels of both responsiveness and demandingness. Given this picture of ideal parenting, psychologists have also come to understand "poor" or impaired parenting as a lack or excess of responsiveness or demandingness. Developmental psychologist Diana Baumrind used this reasoning to create three broad categories to describe parenting styles. Baumrind believed that each different parenting style would correlate with distinct patterns of parent–child interactions and that parenting style could ultimately be used to predict outcomes for children. If research could demonstrate that certain parenting styles benefit children, parents should be advised accordingly. Likewise, if certain parenting behaviors are damaging to children, appropriate interventions and parent training protocol should be implemented.

In 1967, Baumrind launched what is now one of the most famous studies in developmental literature.

Her study began in a nursery school, where she observed children carry out their daily routine for 14 weeks. Based on these observations, children were assigned to one of three categories: energetic-friendly, conflicted-irritable, and impulsive-aggressive. Next, Baumrind interviewed parents and observed them in their interactions with their children at home and in the laboratory setting. She found an association between the three patterns of child behavior and the three parenting styles she had initially outlined.

The first parenting style, *authoritative* parenting, describes parents who are responsive to their children's needs and who also manage to set reasonable expectations and limitations (high responsiveness/high demandingness). Authoritative parents are actively engaged with their children, expressing pleasure and support in their interactions with their children. These parents adhere to core moral standards, address children's misbehavior, and enforce rules calmly and firmly. However, these parents do not carry out rule enforcement in a blind or rigid manner; they are open to, and even solicit, input from the child. This flexibility on the part of the parent helps minimize the child's noncompliance.

Consistent with her hypothesis, Baumrind found that authoritative parents tended to raise the best-adjusted and most mature children. She argued that because authoritative parents balance sensitivity with moderate restrictiveness, their children feel secure in the world and are able to learn and function independently. Further research supports this notion, indicating that children of authoritative parents tend to develop higher levels of self-esteem and self-reliance, value achievement, show greater social competence and cooperation with peers and adults, and demonstrate lower levels of antisocial behavior than other children.

The next category of parenting styles used in the Baumrind study was the *permissive* group, which describes parents who are attuned to their children's needs but fail to set appropriate limits (high responsiveness/low demandingness). Parent–child dyads in this category were fairly warm and loving. However, these parents were lax in setting and enforcing rules and tended to yield to whining or resistance on the part of the children. Unlike authoritative parents, who make their expectations for independent, mature behavior known to their children, few demands are asserted by permissive parents, and children's bad behavior is not challenged.

Children raised by permissive parents often fail to develop age-appropriate self-regulation skills. They tend to be more noncompliant with their parents and other adults, and are less able to recognize and regulate their own emotions and behaviors, leading to aggressive or impulsive actions. These children may appear domineering, aggressive, and uncooperative in interactions with peers. Without the supportive scaffolding provided by authoritative parents, the children of permissive parents tend to show less goal-directed activity and poor self-reliance. Permissive parents do not necessarily see these behaviors as problematic and, in fact, may believe it is unhealthy for children to deny or suppress their urges.

Baumrind's final parenting category, the *authoritarian* style, stands in stark contrast with the permissive style. Where permissive and authoritative parents make efforts to read their children's signals, authoritarian parents are largely insensitive and unresponsive to their children's cues. Authoritarian parents are overcontrolling, setting excessive limits (low responsiveness/high demandingness). Rules are enforced rigidly and harshly by authoritarian parents, and perhaps more troubling, the rationale behind rules is not always made clear to children. Whereas authoritative parents are good at communicating with their children and can thus solicit their children's voluntary cooperation, authoritarian parents may expect obedience for its own sake. Authoritarian parents do not consistently consider their children's opinions.

Children of authoritarian parents tended to fall in the "conflicted-irritable" group. They were more emotionally volatile and sensitive to stressors, and they expressed higher levels of negative affect. It was Baumrind's belief that although these children resent their heavily restricted environments, they are hesitant to assert themselves out of fear of punishment, and this fear leads to feelings of anger, fear, and confusion. Thus, whereas these children may be aggressive at times, they may also rely on passive–aggressive tactics or withdraw from social situations they find stressful. A number of researchers have shown that the harsh punishment and overcontrolling strategies used by some authoritarian parents contribute to externalizing behavior problems in toddlers and preschool children, as well as conduct disorders in adolescence. Thus, just as permissive parenting can lead to problem behavior, so too can power-assertive parenting.

In 1983, Eleanor Maccoby and her colleagues investigated a fourth category of parenting styles to

Baumrind's original three. This last category, known as either the *rejecting*, *neglecting*, or *uninvolved* style, refers to those who are either unable or unwilling to respond to their children's needs. Additionally problematic, these parents do not actively engage in monitoring or rule setting (low responsiveness/low demandingness). Rejecting parents are focused on their own needs rather than those of their children.

Maccoby found that neglected children often develop attachment issues. Others soon proved that these children are also at risk for externalizing problems during preschool. Not only do neglected children miss out on the social training provided by parents, but they are also denied important learning opportunities. This may interrupt cognitive development and academic achievement. Neglected adolescents are at heightened risk for antisocial behavior, including truancy, delinquency, and early sexual activity.

Recent research has further parsed parenting styles into specific parenting behaviors—rather than broad parenting styles—that may contribute to particular clinical disorders in children. These parenting behaviors are understood as falling along a continuum, much like the constructs of high and low responsiveness and demandingness discussed earlier. For example, the research of Jeffrey Wood and his colleagues has shown a clear link between high levels of parental control and childhood anxiety. Parents who excessively manage children's routines and activities, tell their children how to think or feel, make decisions without consulting their child, or who speak or intervene on their child's behalf when such action is unnecessary, are typically considered overcontrolling or intrusive.

Another global construct that has been recently researched is the link between harsh or physical discipline and childhood externalizing disorders. A large body of evidence suggests that harsh punishment, including physical punishments such as spanking and hitting, encourages aggressive and hostile behaviors in children through the process of parental modeling. Not surprisingly, children whose parents are excessively punitive appear to be at higher risk for externalizing problems and conduct disorder.

By identifying the sorts of parenting behaviors that are associated with various childhood disorders, researchers can better understand the mechanisms underlying children's dysfunctions and can incorporate this information into treatment. In other words, if researchers can understand how parents may trigger

or exacerbate children's problem behaviors, treatment can target those specific behaviors.

Limitations and Clinical Applications

Critics of Baumrind and her likeminded colleagues have questioned the true importance of research into parenting style. Some argue that genetic transmission explains why strong parents tend to raise well-adjusted children and impaired parents tend to raise difficult children. Others have argued that it is peer groups, not parents, who leave an indelible print on children's future. Though genetics and social environment are likely important factors in children's socialization, there is strong support indicating that parents play an important role in shaping who their children become.

Longitudinal research suggests that the positive implication of authoritative strategies, and negative implications of permissive, authoritarian, and neglectful styles, remain pertinent for children of all ages. However, it is important to stress that the parent–child relationship is by no means fixed. Through education and intervention, impaired parents can develop new skills and take on a different style of parenting. Furthermore, parent–child relationships are not unidirectional. Parents cannot completely control their dynamics with their children. Children's own genetically determined temperament, for example, can influence the types of parenting strategies used by their caregivers. Some children are born inhibited and reserved and rely greatly on their parents for support. Other children who are more extraverted and independent might perceive parental support as stifling. Additionally, some parents and children are naturally going to be more alike in their temperament than are other dyads. The goodness of fit between parent and child can influence the ease with which parents navigate their relationship with their child and the parenting styles they choose. Siblings bring their own, unique needs and personalities to families, and parents will adjust their approach to parenting accordingly.

It is also important to acknowledge that parenting exists in a cultural context. Research has shown that normative standards for parenting in other cultures vary greatly. In Japanese culture, for example, mothers are extremely responsive to their infants because the Confucian ideal of harmony is highly valued in the parent–child relationship. Whereas Americans might find these parenting patterns "clingy,"

they are typical in Japanese culture. Likewise, Swedish and German cultures encourage early independence; thus, parenting practices common in these nations may seem cold or demanding to American parents. Given these differences in values and normative practices, and in light of the fact that Baumrind's research was based on results from a predominantly middle-class, European American sample, it is reasonable to conclude that Baumrind's predictions regarding parenting styles may not generalize to ethnic or socioeconomic populations outside this norm. In other words, authoritative parenting may not be the "best" option for children from different communities.

Even within the American population, this finding rings true. For example, Beth Kotchick and colleagues found that authoritarian parenting may be more common among African American parents in the United States. Evidence suggests that parents may rely more on authoritarian strategies to protect their children from dangerous, urban environments.

No one style of parenting guarantees a positive or negative outcome for a child. However, research on parenting styles suggest that they are linked with meaningful developmental outcomes in a number of realms—social, academic, and emotional. Though the study of parenting styles originated within developmental psychology, clinical psychologists have adapted parenting theory to explain how maladaptive parenting can lead to negative outcomes for children, and they have used this information to inform intervention and treatment efforts. To expand and hone understanding of effective caregiving, research into parenting styles and child outcomes continues today.

Sarah E. Shea and Lisa W. Coyne

See also Attachment Disorder; Behavior Disorders; Child Abuse; Family Influences; Parenting

Further Readings

Allen, J. P., Hauser, S. T., O'Connor, T. G., Bell, K. L., & Eickholt, C. (1996). The connection of observed hostile family conflict to adolescents' developing autonomy and relatedness with parents. *Development and Psychopathology, 8,* 793–809.

Baumrind, D. (1967). Child care practices anteceding three patterns of preschool behavior. *Genetic Psychology Monographs, 75,* 43–88.

Baumrind, D. (1994). The social context of child maltreatment. *Family Relations, 43,* 360–368.

Belsky, J., Woodworth, S., & Crnic, K. (1996). Trouble in the second year: Three questions about family interaction. *Child Development*, *67*(2), 556–578.

Campbell, S. B. (1994). Hard-to-manage preschool boys: Externalizing behavior, social competence, and family context at two-year follow-up. *Journal of Abnormal Child Psychology*, *22*, 147–166.

Campbell, S. B., Pierce, E. W., Moore, G., Marakovitz, S. (1996). Boys' externalizing problems at elementary school age: Pathways from early behavior problem, maternal control, and family stress. *Development and Psychopathology*, 8,701–719.

Cassidy, J., & Berlin, L. J. (1994). The insecure/ambivalent patterns of attachment: theory and research. *Child Development*, *65*(4), 971–991.

Deater-Deckard, K., Dodge, K. A., Bates, J. E., & Pettit, G. S. (1996). Physical discipline among African American and European American mothers: Links to children's externalizing behaviors. *Developmental Psychology*, *32*(6), 1065–1072.

Kochanksa, G., Aksan, N., & Nichols, K. E. (2003). Maternal power assertion in discipline and moral discourse contexts: Commonalities, differences, and implications for children's moral conduct and cognition. *Developmental Psychology*, *39*(6), 949–963.

Kotchick, B. A., & Forehand, R. (2002). Putting parenting in perspective: A discussion of the contextual factors that shape parenting practices. *Journal of Child and Family Studies*, *11*(3), 255–269.

Lorber, M. F., O'Leary, S. G., & Kendziora, K. T. (2003). Mothers' overactive discipline and their encoding and appraisals of toddler behavior. *Journal of Abnormal Child Psychology*, *31*(5), 485–494.

Maccoby, E. E., & Martin, J. A. (1983). Socialization in the context of the family: Parent-child interactions. In E. M. Hetherington (Ed.), *Socialization, personality, and social development: Vol. 4. Handbook of child psychology.* (Vol. 4, pp. 1–101). New York: Wiley.

Parke, R., & Locke, V. (1999). *Child psychology: A contemporary viewpoint.* Boston: McGraw-Hill.

Patterson, G. R., DeBarshyshe, B., & Ramsey, R. (1989). A developmental perspective on antisocial behavior. *American Psychologist*, *44*, 329–335.

Wood, J. J., McLeod, B. D., Sigman, M., Hwang, W., & Chu, B. C. (2003). Parenting and childhood anxiety: theory, empirical findings, and future directions. *Journal of Child Psychology and Psychiatry*, *44*, 134–151.

PARENT–TEACHER CONFERENCES

Parent–teacher conferences are a formal method of establishing communication between parents and teachers regarding the academic, behavioral, or social strengths and needs of a child. These meetings often provide an initial opportunity to develop a partnership between schools and families. Such partnerships have long been recognized as benefiting the educational and social development of children, as well as promoting positive outcomes for parents and teachers.

Although parent–teacher conferences occur in almost all schools, the manner in which these meetings take place varies considerably. Organizations such as the National Parent Teacher Association and National Association of School Psychologists offer procedural guidelines and identify key components for enhancing the effectiveness of parent–teacher conferences.

Purposes

Parent–teacher conferences can serve several purposes. First, they serve as a vehicle for bidirectional, reciprocal communication between parents and teachers. The intent of the conference is not solely for the school to provide information to the family but for the family to also contribute their expertise and knowledge. Second, conferences facilitate an exchange of information specifically related to the child's development in home and school settings. Third, parent–teacher conferences provide an opportunity to problem-solve potential issues or needs. Fourth, these meetings also serve to outline shared responsibilities between the school and family to further the academic, behavioral, and social development of the child.

Global Assumptions

Betty Jo Simmons identifies three global assumptions that are essential to producing effective parent–teacher conferences:

1. *Assume good will.* This implies the belief that both parents and teachers are coming in to the conference with positive motivations and positive results in mind.

2. *Assume competence.* Maintain the viewpoint that parents and teachers share expertise about a child, in different arenas. In an effective conference, both parents and teachers are able to share their expert opinions about a child and his or her functioning.

3. *Assume shared responsibility.* Both parents and teachers are invested in the life and development of

a student. Both parties must agree to cooperate and work together to seek a common solution that will lead to a child's success.

General Guidelines

There are several general guidelines that enhance the effectiveness of parent–teacher conferences. First, it is helpful to keep the meeting solution focused, following a problem-solving framework. This prevents the meeting from concentrating on what is wrong and redirects the focus to how the issue might be addressed. It also provides the conference with purposeful intent. Second, strategies developed in the conference should be strength-based. Concerns should be addressed in ways that will build upon the strengths and resources possessed by the student, family, and teacher. Third, the discussion should be centered on the child's performance. This separates the performance of a behavior from the entity of the child. Thus, it is not that the child is "bad"; rather, it is that the behavior is not desired. Fourth, it is beneficial to develop a shared ownership of the problem solution. This enhances everyone's commitment, increases the fidelity of intervention developed, and allows for a sharing of resources and demands required in implementing the intervention strategy. Fifth, goals developed should be objective and meaningful. This allows for the evaluation of the effectiveness of the intervention strategy and ensures that the goals are realistic and achievable. Finally, a system for future communication between parents and teacher should be developed and implemented. This provides follow-up information on current issues, promotes future collaboration, and prevents future problems from occurring.

Parent-Specific Considerations

Parent–teacher conferences are more effective when parents approach the conference as an opportunity to establish a cooperative partnership in which both partners are working for the best interests of the child. Andrea Canter provides steps that parents can take to promote a positive and productive conference.

- *Prepare and review appropriate materials.* Parents should gather and review materials (i.e., health records, previous school records) that could provide relevant information for the child's teacher.
- *Talk with your child before the conference takes place.* Parents should communicate with their child the importance of the meeting and what will be discussed and find out what concerns the child has related to academics or social issues.
- *Develop a list of questions for the teacher.* These questions should be related to the child's academic, behavioral, and social performance and should focus on both strengths and needs.
- *Be ready to collaborate.* Parents should provide their expertise and opinions about the topics discussed and ask questions for clarification.

Teacher-Specific Considerations

Although the conference should be a collaborative experience, teachers often set the tone for the meeting. Simmons outlines the following teacher-oriented considerations to enhance the conference's effectiveness and productivity.

- *Consider parents' physical comforts.* Teachers should provide a comfortable and inviting setting for parents (i.e., offer refreshments, provide adult-size chairs).
- *Seek workable solutions.* Teachers should invite parents to share their opinions, value their perspectives, and support parents' endeavors to assist their child.
- *Listen empathically and with understanding.* Using reflective statements and paraphrasing to summarize what has been discussed helps ensure understanding.
- *Offer information and resources.* Teachers should provide parents with additional information or resources if desired (i.e., referral sources, Web sites, pamphlets, parent support groups).
- *Avoid jargon.* When discussing the student's performance and progress, teachers should use terminology that parents understand. Effective teachers also clarify questions as they arise.
- *Take a positive focus.* Maintaining a solution-focused orientation can shift emphasis from a problem-based approach to a strength-based perspective. This can reduce the emphasis from problem areas and help promote in parents a sense of pride in their child.
- *Outline shared responsibilities.* Providing a clear summary of expectations and responsibilities for all parties involved in the conference promotes a sense of direction and focus. It also facilities shared ownership of the child's needs and strategies developed through the conference.

John Warren Eagle and Jessica Oeth

See also Early Intervention Programs; Home Environment and Academic Intrinsic Motivation; Individualized Education Program; Mainstreaming

Further Readings

Adams, L. D. (2000). Conferring with Amanda's mom. *Childhood Education*, *76*, 106–107.

Canter, A. (2004). *Parent–teacher conferences: A guide for parents.* Bethesda, MD: National Association of School Psychologists.

Christenson, S. L., & Sheridan, S. M. (2001). *Schools and families: Creating essential connections for learning.* New York: Guilford Press.

Epstein, J. L. (1995). School/family/community partnerships: Caring for the children we share. *Phi Delta Kappan*, *76*, 701–712.

Simmons, B. J. (2002). Facilitative conferences: Parents and teachers working together. *The Clearing House*, *76*, 88–93.

PEER-ASSISTED LEARNING

Peer-assisted learning can be defined as the acquisition of knowledge and skill through active helping and supporting among status equals or matched companions. It involves people from similar social groupings who are not professional teachers helping each other to learn and learning themselves by so doing.

Where much older helpers work with much younger learners, the differential in levels of ability and interest can be understimulating for the helpers, who are unlikely to gain cognitively. The peer helping interaction is different from that between a professional teacher and a child or young person, and it has different advantages and disadvantages. Consequently, there has been more interest in deploying helpers whose capabilities are nearer to those of the helped. The helper is intended to be "learning by teaching" and also be a more proximate and credible model.

In this entry, peer learning is defined, and a typology and effects are considered. The longest established and most intensively researched forms of peer learning are peer tutoring and cooperative learning.

Peer tutoring (PT) is characterized by specific role taking as tutor or tutee, with high focus on curriculum content and usually also on clear procedures for interaction, in which participants receive generic and/or specific training. Some peer tutoring methods scaffold the interaction with structured materials, whereas others prescribe structured interactive behaviors that can be effectively applied to any materials of interest.

Cooperative learning (CL) has been described as "structuring positive interdependence" in pursuit of a specific shared goal or output. This is likely to involve the specification of goals, tasks, resources, roles, and rewards by the teacher, who facilitates or more firmly guides the interactive process. Typically operated in small groups of about six heterogeneous learners, CL requires previous training to ensure equal participation and simultaneous interaction, synergy, and added value.

One of the most important changes over time has been a greater focus on implementation integrity. This has involved sharpening awareness of the organizational variables in the delivery of peer learning, such as the following:

- *Curriculum content*—knowledge or skills or combination to be covered.
- *Contact constellation*—the size of group can vary from 2 to 30 or more. More intensive is PL in pairs.
- *Within or between institutions*—although most PL takes place within the same institution, it can also take place between different institutions.
- *Year of study*—helpers and helped may be from the same or different years of study and/or be the same or different ages.
- *Ability*—although many projects operate cross-ability (even if they are same-age/year), there is increasing interest in same-ability.
- *Role continuity*—especially in same-ability projects, structured switching of roles at strategic moments (reciprocal PL) can have the advantage of involving greater novelty and a wider boost to self-esteem.
- *Time*—in regular class contact time, outside of this, or in a combination of both, depending on the extent to which it is substitutional or supplementary for regular teaching.
- *Helper characteristics*—if helpers are those who are merely average (or even less), all partners should find some challenge in their joint activities. Although the gain of the helped might not be so great, the aggregate gain of both combined may be greater.
- *Characteristics of the helped*—projects may be for all or a targeted subgroup.
- *Objectives*—projects may target intellectual (cognitive) gains, formal academic achievement, affective and attitudinal gains, social and emotional gains, self-image and self-concept gains, or any combination.
- *Voluntary or compulsory*—some projects require participation, whereas in others, helpers select the participants.
- *Reinforcement*—some projects involve extrinsic reinforcement for the helpers (and sometimes also the helped), whereas others rely on intrinsic motivation.

The following aspects of organization need to be considered: context, objectives, curriculum area, participants, helping technique, contact, materials, training, process monitoring, assessment of students, and evaluation and feedback.

When PT or CL are implemented with reasonably high implementation integrity, results are typically very good in the cognitive area for all partners. Both CL and PT can simultaneously yield gains in transferable social and communication skills and in affective functioning (improvements in self-esteem, liking for partner or subject area). Peer counseling can be at least as effective as adult counseling; the same is true for "peer education" (peers offering credible and reliable information about sensitive life issues and the opportunity to discuss these issues in an informal peer group setting). Peer monitoring has been extended beyond unwanted behavior to learning behavior and operated on a classwide basis with excellent results. Peer assessment (peers evaluating the products or outcomes of learning of others in the group) is widely used. Having learners "mark," "grade," or quantitatively assess the products of their peers can result in learner social discomfort. More cognitively demanding for the assessor is giving formative and qualitative feedback, which is likely to be more socially comfortable and more useful to the assessee.

In recent years learners who themselves have educational challenges have been found to act effectively as tutors to other learners. Gains for the tutors themselves have been increasingly emphasized. Peer learning has been noted to be among the most cost-effective of learning strategies.

Keith Topping

See also Competition; Constructivism; Cooperative Learning; Learning Style; Scaffolding; Self-Efficacy; Social Learning Theory; Zone of Proximal Development

Further Readings

Cohen, P. A., Kulik, J. A., & Kulik, C. C. (1982). Educational outcomes of tutoring: A meta-analysis of findings. *American Educational Research Journal*, *19*(2), 237–248.

Johnson, D. W., & Johnson, R. T. (1999). *Learning together and alone.* Englewood Cliffs, NJ: Prentice Hall.

Rohrbeck, C. A., Ginsburg-Block, M. D., Fantuzzo, J. W., & Miller, T. R. (2003). Peer-assisted learning interventions with elementary school students. *Journal of Educational Psychology*, *95*(2), 240–257.

Topping, K. J. (2001). *Thinking reading writing.* New York: Continuum International.

PEER INFLUENCES

Humans are inherently social beings who interact with one another over the course of life. From these interactions, people change and grow over time. The term *peer influences* refers to the body of research aimed at understanding how people change and grow from interacting with others who are of the same status or age. Popular culture has coined terms like *peer pressure* to describe the influence of peers on child development. Educational psychology has been interested in how peers influence learning and behavioral changes in school performance. Consequently, much of the research has focused on the peer influences of school-age children. This entry discusses the peer influences knowledge base broadly and then focuses on two areas that are most important to education and schooling: deviant behavior and prosocial behavior.

Understanding Peer Influences

It is important to understand peer influences within the broader area of peer relations research and differentiate between the two. Several important peer relations concepts are discussed first. Then, several important peer influences concepts are discussed.

General Concepts of Peer Relations

The peer influences knowledge base is part of a larger research area called *peer relations*, which refers to the study of children's social world of peers, how they interact and form relationships, how they change over time, and how they affect children's development. For example, peer relations research examines how children make friends, how children's friendships change over time, and how having friends affects children's development of social competence. Therefore, research on peer influences overlaps with a rather large portion of the peer relations research that examines how peers affect children's development in general. However, this entry focuses on the narrower and more recent area of peer influences

research that examines how children's behavior changes and grows in positive and negative ways from interacting with their peers over time (e.g., how friends negatively influence one another to smoke).

Still, because it is important to understand peer influences within this larger context of social development, several general themes of children's peer relations need to be discussed. First, children's peer relations systematically change in form and function as children grow. Second, peer relations are powerful in shaping children's development. In preschool and early childhood, peer relations do exist, but they are ephemeral and smaller in size and are largely affected by adult choices (e.g., play groups, recess). From middle childhood through adolescence, children interact with their peers more frequently than before, peer groups tend to be larger with much less adult supervision, and peer interactions occur in more settings (e.g., outside home and school). Children also spend their time talking, playing organized games or sports, and "hanging out" more as they get older. Therefore, youth spend more time with their peers and are more influenced by them during this time than ever before.

Peer relations have been studied most commonly according to social configurations that are categorized within two broad types: *dyads* and *groups*. Dyads include best friends, friends, enemies, and romantic relationships. Groups include smaller peer groups of friends who hang out together (e.g., social networks, cliques, crowds, counseling groups) and larger groups such as classrooms, which typically consist of dyads and several peer groups.

General Concepts of Peer Influences

Peer influences can occur within any of the peer relations configurations. Typically, children change their behavior (target) because they are influenced to change by another peer (influencer). Importantly, targets and influencers may also be more than one child.

Change in the target is assumed to occur through socialization or from contact and interactions with a peer influencer. However, it is also possible that change is not really occurring but seems to have happened because the target child *selects* friends who are similar to the target child already. For example, a boy who is nice hangs out with other nice boys because they are alike (or nice); thus, the boy was already nice and had not been influenced by his peers to become

nice at all. At this time, most authorities on peer influences agree that both socialization and selection appear to be working simultaneously. Still, socialization effects are real and help explain how peer influences occur. Within a socialization perspective, there are more specific explanations of how peer influences occur, such as reinforcement, identification with peers, social norms, identity enhancement, and regulation of social roles and dominance.

The effects of peer influences and how they work more specifically depends on a variety of factors. First, peer influences depend on the behavioral domain or type of behavior, such as delinquent activity, sexual activity, academic performance, and so on. Second, peer influences will vary depending on the developmental age of the child and peers. Third, peer influences will depend on the characteristics of the influencing peers, the target child, and their relationships with others. Finally, broader contextual factors like the size of the peer group are also important to consider. To illustrate these factors, a girl may be more or less influenced depending on whether her peers are influencing her to do drugs or study hard, whether she is 7 years old or 16 years old, whether she is an ethnic minority and her peers are ethnic minorities, and whether she is a friend or an enemy to her peers.

In sum, understanding peer influences is a complex task. There are a variety of factors to consider that affect how, and whether, peers influence each other. In the following sections, peer influences are discussed in relation to a specific behavioral domain; a discussion of the factors that are known to affect peer influence for that specific behavioral domain is included.

Deviant Peer Influence

The most popular and perhaps the most crucial area of peer influence research focuses on the development of deviant behavior. Affiliation with a deviant peer group has a strong impact on the development of delinquent behaviors (e.g., minor acts, severe offending, and substance abuse) and, even after accounting for previous delinquency, predicts later delinquency. The onset of problem behaviors (e.g., stealing, selling drugs, attacking someone with a weapon) and dropping out of school are also linked to having deviant friends.

Deviant peer groups are a primary contributor to a variety of antisocial behaviors during adolescence. There are two main groups of antisocial children: *early starters* and *late starters*. Early starters experience a range of psychosocial deficits (neurocognitive, familial, and social) that contribute to the early- and long-term development of antisocial behavior. Late starters do not engage in antisocial behavior until adolescence, when they consider it normal to experiment with delinquent behavior. As adolescents mature into adulthood, they stop engaging in antisocial behaviors. In an illustration of deviant peer group influences on late starters, youth who were members of a deviant peer group in the eighth grade had significantly more encounters with the police compared with youth who were not members of a deviant peer group. For early starters, deviant peer groups affect the severity of delinquent acts they commit rather than affecting whether they will engage in them.

Teen girls are similar in that they are influenced by deviant peers, but they are more likely to engage in early sexual activity, substance use, and covert antisocial activity. Additionally, girls appear to be influenced by older antisocial boys. Those most susceptible to the influence of deviant peers are girls who enter puberty early, have learning difficulties, and have depressed moods.

Deviant Peer Contagion

The work of Thomas Dishion, Kenneth Dodge, and their colleagues has shed light on the specific problem of *deviant peer contagion*, which they refer to as the unintended negative side effects that occur from grouping peers in an intervention program designed to be therapeutic, educational, or provide other helping outcomes. This phenomenon was first observed by Dishion and his colleagues during a therapeutic group intervention for deviant and delinquent behaviors, which aggregated youth who exhibited deviant behaviors together in a treatment group. They found that when youth were grouped together for an intervention, their deviant behaviors (i.e., tobacco use and externalizing behaviors) increased compared to those of the control group; furthermore, these effects persisted over a 3-year follow-up period. They found that children provided reinforcement for deviant behavior by laughing and giving attention to those who talked about antisocial activities. More recently, Dishion's research team have theorized that the deviant talk with peers results in having a "connection with one another," which may be the key to increased motivation to engage in delinquency.

The phenomena of deviant peer contagion and, more generally, deviant peer influences, is troublesome, given that the aggregation of children and youth into groups is a common (and often preferred) practice in education, mental health programs, juvenile justice interventions, and community programs. In education, for example, children are grouped together for certain special education services (academic interventions); disciplinary reasons (after-school detention, alternative schools); or psychoeducational and therapeutic reasons (social skills or self-esteem groups). In mental health, group counseling and therapy is frequently used as an intervention method for children and youth. The juvenile justice system places delinquent youth into juvenile detention centers. Finally, although less common and less understood, community programs sometimes have youth convene in a community center without much adult supervision.

Dishion and Dodge recommend using an ecological framework to understand deviant peer influence and contagion. The ecological framework they propose contains three levels: individual, program or intervention service, and context. The *individual level* refers to characteristics, such as age, gender, history of deviant behavior, and so on, of the youth who is participating in the intervention or program. The *program* or *intervention service level* refers to the formal and informal social interactions among the participants and adult leader that occur during the intervention and as a result of the intervention (e.g., number of delinquent and nondelinquent participants, number of leaders). Finally, the *context level* is the setting with which the intervention or program is integrated. The settings are community, public school, mental health, and juvenile justice. To comprehensively understand deviant peer influence all three of these levels and their interaction should be considered.

Because educational psychology is primarily concerned with education and schooling, specific research findings regarding deviant peer influence are discussed within this area. Schools are obvious settings in which to study deviant peer influence because they aggregate children and youth naturally as well as in with more purpose. Examples include grouping by achievement or ability level (e.g., differentiated learning groups, gifted, special education services); general classroom assignments; the assignment of students with behavioral

difficulties to disciplinary rooms or alternative schools; and other educational practices (e.g., retention, suspension, and expulsion). It is also important to note that intervention programs are increasingly being implemented in schools that group students who are exhibiting behavioral difficulties. Research findings have documented negative long-term effects of the natural aggregation of aggressive students in the classroom setting. Several researchers have found that classrooms containing a high degree of class aggression negatively influenced the individual behavior of students currently and in the future. Still, there is not enough research evidence to know whether certain educational practices lead to deviant peer influence.

Positive Peer Influence

Although much of the research attention has been on deviant peer influence, researchers should not neglect the phenomenon of positive peer influence. Indeed, peers can influence one another in positive and prosocial ways as well. Peers can influence each other's prosocial behaviors simply by being nearby during times of need. *Prosocial behaviors* are positive responses intended to help or benefit another person or group and are often intended to put another person in a more positive state. Examples of prosocial behaviors in children include standing up to bullies, helping with schoolwork, and asking children to play. Several theorists suggest that prosocial responding is motivated by empathy.

Empathy is a set of emotional and cognitive constructs that allows one to experience the state of another. When one witnesses another in distress, there is first an automatic empathic reaction. After processing the information, one's reaction can then turn into sympathy or personal distress. Sympathy is a general feeling of concern for another person because of his or her situation. Sympathy and empathy motivate one to help another person. These two terms are often used interchangeably and not always differentiated in the literature because the outcome of helping is similar. Conversely, the initial empathic reaction could lead to personal distress.

Personal distress is characterized by an aversive emotional reaction to the situation. The focus changes from the other's situation to one's own distress. The behavioral outcome could be helping, ignoring, or escaping; it depends on whatever is the easiest way to relieve one's own distress. Experimental research with

adults has found a strong link between empathy and prosocial behavior; the connection is not as strong in the research with children although it appears to increase with age. One theory is that empathy develops concurrently with one's sense of self and others. Thus, it is suggested that peers influence prosocial behavior by eliciting empathic responses in others.

Peers also influence prosocial behavior through modeling and social learning among both young children and adolescents. For example, withdrawn preschool and kindergarten children become more sociable and prosocial after regular play sessions with other children. Prosocial behavior in kindergarten classrooms also increases when children are encouraged to report cooperative or friendly behaviors to the teacher. In addition, several studies suggest that children who behave prosocially are also more likely to receive prosocial behaviors from their classmates. In fact, being the recipient of prosocial behaviors may be the best predictor of subsequent prosocial behaviors. This suggests that young children respond positively to prosocial behaviors from peers, and are likely to imitate them.

Close friendships are also likely to influence prosocial behavior and goals in adolescents. Through self- and peer-reporting, adolescents' prosocial behaviors were found to be highly related to those of their best friends; this relation increased over time. Group processes and norms also moderate behaviors; classroom motivation varies by peer group approval or disapproval. Overall, the research suggests that prosocial behaviors elicit prosocial responses. In addition, prosocial behaviors easily generalize; prosocial children are likely to give and receive a wide variety of positive behaviors. This builds a positive, happy, and safe classroom environment where children and teachers can focus on academics.

Future Considerations

Peers are clearly powerful influences on human behavior. A full understanding of peer influence requires general mastery of the peer relations literature in addition to knowledge of the more specific literature on deviant peer influence and positive peer influence. Educators and psychologists need to understand the complex phenomenon of peer influences through systematic research and evaluation of current educational practices and policy. Without well-developed knowledge of peer influences, schools may inadvertently

encourage children's deviant behavior and discourage their prosocial behavior.

Samuel Y. Song and Natalie M. Siegel

See also Aggression; Conduct Disorders; Cooperative Learning; Friendship; Gangs; Peer-Assisted Learning; Social Development

Further Readings

Dishion, T. J., McCord, J., & Poulin, F. (1999). When interventions harm: Peer groups and problem behavior. *American Psychologist, 54*, 755–764.

Dodge, K. A., Dishion, T. J., & Lansford, J. E. (2006). *Deviant peer influences in programs for youth: Problems and solutions.* New York: Guilford Press.

Eisenberg, N., & Miller, P. A. (1987). The relationship of empathy to prosocial and related behaviors. *Psychological Bulletin, 101*(1), 91–119.

Hartup, W. W. (2005). Peer interaction: What causes what? *Journal of Abnormal Child Psychology, 33*(3), 387–394.

Iannotti, R. J. (1985). Naturalistic and structured assessments of prosocial behavior in preschool children: The influence of empathy and perspective taking. *Developmental Psychology, 21*(1), 46–55.

Kupersmidt, J. B., & Dodge, K. A. (2004). *Children's peer relations: From development to intervention.* Washington, DC: American Psychological Association.

Staub, E. (2003). *The psychology of good and evil: Why children, adults, and groups help and harm others.* Cambridge, UK: Cambridge University Press.

PERCENTILE RANK

There are many ways to measure and evaluate things, but one of the most common is by comparison. *Percentile ranks* indicate the percentage of scores lower than any given value on a scale from 1–99. For example, if 100 students took a history test, a percentile rank of 36 would mean that the student's score was better than that of 36 other students. If 200 students had taken the test, a percentile rank of 36 would indicate that the student had scored higher than 72 others.

Percentile ranks are based on *cumulative frequencies*, which are the number of scores equal to or less than a given value. Cumulative frequencies are found by placing all of the scores in order from lowest to highest and adding the number of scores (i.e., *frequency*) for each score or range of scores beginning with the smallest. To illustrate, Table 1 (which is the basis for the example illustrating the procedure

outlined below) shows a simple frequency distribution (a display of the number of occurrences of each of the scores) for 50 people on a test with 20 questions. In the first three columns, X indicates the value of the scores, f is the frequency of each score, and cf is the cumulative frequency. Thus, you can see that 3 people answered 8 questions correctly, and a total of 18 people answered 8 or fewer questions correctly (i.e., the cumulative frequency, cf, for $X = 8$ is 18).

Calculation of Percentile Ranks in Simple Frequency Distributions

Calculation of percentile ranks (PR) in a simple frequency distribution is based on the assumption that

Table 1 Percentile Ranks for Scores in a Simple Frequency Distribution

X	f	cf	cf_{mp}	cf_{mp}/F	PR
20	1	50	49.5	0.99	99
19	1	49	48.5	0.97	97
18	2	48	47	0.94	94
17	0	46	46	0.92	92
16	1	46	45.5	0.91	91
15	0	45	45	0.90	90
14	3	45	43.5	0.87	87
13	2	42	41	0.82	82
12	5	40	37.5	0.75	75
11	7	35	31.5	0.63	63
10	6	28	25	0.50	50
9	4	22	20	0.40	40
8	3	18	16.5	0.33	33
7	5	15	12.5	0.25	25
6	2	10	9	0.18	18
5	4	8	6	0.12	12
4	2	4	3	0.06	6
3	1	2	1.5	0.03	3
2	0	1	1	0.02	2
1	1	1	0.5	0.01	1
0	0	0	0	0	1

each score is the midpoint of a range of possible scores (e.g., for $X = 8$, the interval from 7.5 to 8.5). Finding the percentile rank of a score thus involves finding the cumulative frequency for the midpoint of the interval (cf_{mp}) corresponding to the score of interest and involves the following steps:

1. Find the cumulative frequency for the midpoint of the interval (cf_{mp}) of the score of interest by averaging the cumulative frequency of the score with the cumulative frequency of the score below it. For example, for $X = 8$ in Table 1,

$$cf_{mp} = \frac{18 + 15}{2} = \frac{33}{2} = 16.5.$$

2. Divide the cf_{mp} by the total frequency (F) of the distribution. Thus, for $X = 8$,

$$\frac{cf_{mp}}{N} = \frac{16.5}{50} = .33.$$

3. Multiply the resulting proportion by 100 and round to the nearest whole number to find the percentile rank (e.g., $.33 \times 100 = 33$).

For ease of computation, the formula that summarizes this procedure (in which N is the total number of scores) can be simplified as follows:

$$PR_X = \frac{\frac{cf + cf_{X-1}}{2}}{N} \times 100 = \frac{cf_X + cf_{X-1}}{2N} \times 100$$
$$= \frac{cf_X + cf_{X-1}}{N} \times 50.$$

In Table 1, note that the percentile ranks of scores of both 0 and 1 are shown as 1, even though no one received a score lower than 1. This is true because by convention, percentile ranks are reported on a scale from 1–99.

Calculation of Percentile Ranks in Grouped Frequency Distributions

When frequencies are reported in tables with intervals greater than one (sometimes referred to as *grouped frequency distributions*), percentile ranks are estimated based on the assumption that the scores are evenly distributed throughout the interval. For example, Table 2 is an excerpt from a frequency distribution of 200 scores on a 100-point scale. Although the intervals typically are depicted using whole numbers

(e.g., 61–65), this is shorthand for intervals that extend beyond the values shown to their *real limits* (i.e., 60.5–65.5). The computational formula for percentile rank of a score (PR_X) is

$$PR_X = \frac{cf_{ll} + \left(\frac{X - X_{ll}}{i}\right)f_i}{N} \times 100,$$

where X is the score of interest, cf_{ll} is the cumulative frequency of the lower limit of the interval in which it occurs, X_{ll} is the lower real limit of that interval, i is the width of the interval, and f_i is the frequency of the interval. It can be noted that the value of cf_{ll} for a score is the same as the cumulative frequency of the interval below which it occurs. Thus, based on the data presented in Table 2, the percentile rank of someone who received a score of 64 would be

$$PR_{64} = \frac{cf_{ll} + \left(\frac{X - X_{ll}}{i}\right)f_i}{N} \times 100$$
$$= \frac{135 + \left(\frac{64 - 60.5}{5}\right)12}{200} \times 100$$
$$= \frac{135 + \left(\frac{3.5}{5} \times 12\right)}{200} \times 100$$
$$= \frac{135 + (.7 \times 12)}{200} \times 100$$
$$= \frac{135 + 8.4}{200} \times 100$$
$$= \frac{135 + 8.4}{200} \times 100$$
$$= \frac{143.4}{200} \times 100$$
$$= .717 \times 100$$
$$= 72.$$

Percentile Ranks as Milestones in Distributions

Three percentile ranks are sufficiently important and commonly used that they have acquired special names. The 50th percentile is referred to as the *median*. Because the median is the score that falls exactly in the middle of all scores, with half of the scores lower and half higher, it is one of three commonly used measures of the *central tendency* of a distribution, along with the *mean* (numerical average) and the *mode* (the

Table 2   Percentile Ranks for Scores in a Grouped Frequency Distribution

X	Real Limits	f	cf
96–100	95.5–100.5	5	200
91–95	90.5–95.5	4	195
86–90	85.5–90.5	6	191
–	–	–	–
–	–	–	–
–	–	–	–
66–70	65.5–70.5	10	157
61–65	60.5–65.5	12	147
56–60	55.5–60.5	9	135
–	–	–	–
–	–	–	–
–	–	–	–

score with the highest frequency). In some cases, such as *skewed* distributions (asymmetrical distributions with scores piled up at one end of the scale and tailing off toward the other), the median is considered to provide the most appropriate measure of central tendency. For example, house prices and annual household incomes have *positively skewed* distributions, with most at the low end of the scale and a few that run to astronomical values (e.g., Michael Jordan, Oprah Winfrey, Bill Gates). These atypical values have a disproportionate effect on the mean, making it misleading as a measure of the typical value. In such cases, the median is the preferred measure of central tendency.

The 25th percentile is referred to as the *first quartile*, because a quarter of the scores are lower than it. Similarly, the 75th percentile is the *third quartile*. These values are sometimes used to describe the *variability*, or spread, of a set of scores. The *semi-interquartile range* (*SEQR*) is half the distance between the first and third quartiles. For example, in Table 1,

$$SEQR = \frac{Q_3 - Q_1}{2} = \frac{12 - 7}{2} = \frac{5}{2} = 2.5.$$

Because the semi-interquartile range measures the spread of the middle 50% of the scores, it is less influenced by extreme scores than are other measures of variability, such as the *range* or *standard deviation*.

Use of percentile ranks in reporting and interpreting data has many advantages, including ease of understanding. A disadvantage of percentile ranks, however, is that unlike *standard scores* (which provide a means of interpreting a score in terms of its distance to the average score that takes into account the overall variability or spread of the set of scores), intervals between percentile ranks do not have to be equal. For example, in Table 1, the difference between scores at the 87th and 97th percentile ($19 - 14 = 5$) is larger than the difference between the 40th and 50th percentile ($10 - 9 = 1$) despite the fact that the difference between the two percentile ranks is 10 in each case.

Ernest T. Goetz

See also Descriptive Statistics; Measurement; Median; Norm-Referenced Tests; Standard Scores

Further Readings

Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice Hall.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Harcourt Brace Jovanovich.

Lyman, H. B. (1998). *Test scores and what they mean* (6th ed.). Boston: Allyn & Bacon.

Shavelson, R. J. (1996). *Statistical reasoning for the behavioral sciences* (3rd ed.). Boston: Allyn & Bacon.

PERCEPTUAL DEVELOPMENT

Sensations are units of information received from the environment, such as a visual feature or a sound. *Perceptions* are organized and interpreted sensations, such as recognizing a face or interpreting a sequence of sounds as a familiar song. These concepts represent more of a historical distinction than a functional distinction; nevertheless, the distinction between sensation and perception continues to be made by researchers and textbook writers. *Perceptual development* is the emerging capacity to detect information from the environment and from internal sources to adapt to and function within the world. Knowledge about the perceptual development of infants has expanded more rapidly than that of older children. This discrepancy has occurred because the earliest appearance of various perceptual capacities has been emphasized. In

addition, knowledge about visual development has expanded more rapidly than knowledge about the development of other perceptual systems. This discrepancy has occurred, in part, because of the belief that the visual system provides the best source of external information.

In educational psychology, a quick review of several introductory textbooks reveals that little attention is given to sensation and perception. Most textbooks mention information processing theories of knowledge acquisition and include a three component model of human memory consisting of sensory registers, short-term working memory, and long-term memory. Within this model, the sensory registers receive environmental input. These registers have a large capacity to receive information, but the information quickly fades away unless it is transferred to short-term working memory. The other section of educational psychology textbooks in which sensations and perceptions are mentioned concerns students with special needs, in particular, those with a sensory challenge such as a visual or auditory impairment. Despite this limited coverage, much knowledge has been obtained on perceptual development within the last half-century and a familiarity with this knowledge could facilitate the teaching of students.

This entry is organized in the following way. The development of the five sensory systems that focus on external information and on two sensory systems that provide "internal" information, is presented first. Next is a section on the coordination of information from multiple sensory sources and on the coordination of the perceptual and motor systems. Finally, there is a section on brain development and on the role of experience on perceptual development.

Development of Individual Senses

Vision

In 1890, William James described the newborn infant's visual world as a blooming buzzing confusion. This view represents a strong empiricist perspective that perception develops through learning. It is now known that the newborn's view of the world is not this confusing and also that many important perceptual developments occur within the first year of life. Some principles of looking in infancy include opening the eyes when the light is not too bright, making broad eye movements until an edge (an area

of high contrast) is found, and then continuing to look in the general vicinity of the edge (while making eye movements across the edge). These principles maximize the firing rate of neurons in the eye and brain, which may facilitate further visual developments. Several functions of vision are presented in the following paragraphs, including kinds of eye movements, pattern detection, image features, and depth perception.

Two eye movement types used to localize objects in the environment include saccades and smooth pursuit. Saccadic eye movements shift the eyes from one position to another. Early in infancy, multiple saccades may be required to get to a distant target. Among adults, a saccade will typically bring the eye at least 90% of the distance to the target. Smooth pursuit allows one to maintain fixation on a moving object. Infants cannot use smooth pursuit to track moving objects in the first month of life but, by 3 months of age, smooth pursuit eye movements are common.

Pattern detection refers to the ability to perceive whether an image has contours or edges. Visual acuity is one of the most basic measures of pattern detection, and it is used to determine the maximum resolving capacity of the eyes. An eye chart is used for letter recognition, and a ratio is reported indicating the distance at which a line of letters can be correctly seen to the farthest distance at which a person with "normal vision" can read the same line. Hence, a ratio of 20/20 indicates that the viewer correctly recognizes letters at 20 feet that one with "normal vision" recognizes at 20 feet. Likewise, a ratio of 20/30 indicates that the viewer recognizes at 20 feet what the average person recognizes at 30 feet. Hence, the person's acuity is poorer than average. Infants cannot recognize letters, but other techniques have been developed to measure their acuity. In the newborn, acuity is 20/600 or worse, but it rapidly improves to about 20/200 by 4 months, 20/100 at 6 months, and 20/50 at one year. One may be "legally blind" with visual acuity of 20/200 or less; however, even very young infants can see nearby people and objects.

Image features that attract an infant's attention have also been examined. In one technique, different pictures are shown to an infant, and the infant is observed to determine which picture he or she looks at for a longer time. Infants look longer at high-contrast images such as stripes and bull's-eyes than at solid-colored objects. They also look at faces for a long time. Facial patterns may receive special attention from infants

because of the survival value of connecting with other people.

Depth perception is the process of determining the distance of objects, and it is a crucial ability for determining the spatial layout of the environment. Three sources of information are potentially available for determining depth: binocular information, static monocular information, and dynamic (or kinetic) information. Infants seem to be sensitive to at least some aspects of each of these sources of information in the first year of life.

Two binocular cues to depth are vergence and disparity. Vergence eye movements involve moving the eyes toward or away from each other in the horizontal plane, and this ability enables the eyes to focus on objects of different distances. Binocular disparity is the slightly different image that the two eyes receive because they are set apart by several centimeters. This information can be used to detect the difference in depth between two objects. Humans are very sensitive to this depth information.

Static monocular cues, or pictorial cues, are often used by artists to create the impression of depth on a flat surface. These cues include interposition, in which one object partly conceals another; aerial perspective, in which far away objects are less clear than close objects; linear perspective, in which the size and space of more distant objects decreases; and relative size, which occurs when two identical images of different sizes are viewed and the larger image appears to be closer to the viewer. Sensitivity to some static monocular depth cues has been found by 7 months of age.

Dynamic or kinetic information may also be obtained with only one eye, but this information cannot be represented in a two-dimensional image. Kinetic information is produced by changes in the retinal image over time. These changes can be created by motion of the object in the environment, for example, the expansion of the retinal image in an approaching object. These changes can also be created by movement of the observer. For example, even very slight lateral head movements may create motion parallax in which objects in the visual field that are closer than the fixation point move rapidly in the opposite direction of the observer's head movement whereas objects beyond the fixation point move slowly in the same direction as the head movement.

Despite the rapid development of the visual system during infancy as indicated in several capabilities, visual perception continues to develop and become more efficient throughout childhood. For example, visual skills become better focused, better organized, and more confined to meaningful environmental features as children age.

Audition

Like the visual system, the auditory system is also functional at birth and is fast developing. Fetuses hear before birth: They react to sounds by moving around, and their heart rate increases. Newborns recognize stories that their mothers repeatedly read out loud before they were born, and they prefer to listen to their mothers' voices over those of stranger. Newborns also prefer infant-directed speech over adult-directed speech. Infant-directed speech (*motherese, parentese*) is high in pitch, is slow, and has exaggerated rises and falls. This speech may help infants with language acquisition by directing their attention to particular words.

Infants are sensitive to sounds that seem to help them learn about people. They prefer sounds in the same frequency ranges in which speech occurs and that cover a range of frequencies (like speech) over pure tones; they prefer the language to which they are exposed over other languages; and by 4 months they attend to their own name.

Infants are also sensitive to many contrasting phonemes (speech sounds), such as the distinction between *pa* and *ba*, including phonemes that are not meaningful in their native language. During the first year, they lose sensitivity to many of the contrasts that are not used in the language to which they are exposed. Hence, their ability to discriminate sounds not meaningful in their native language decreases as their sensitivity to sound patterns within their native language continues to improve.

Auditory perception continues to develop during childhood. For example, the ability to detect low-frequency tones develops over several years.

Smell, Taste, and Touch

Researchers have given less attention to the senses of smell, taste, and touch than they have to vision and hearing. These senses are functional early on, and they may be adaptive to the survival of the infant. For example, within the first week of life babies recognize and turn to the smell of their own mother's breast pad

over that of another woman. Infants also prefer familiar odors over ones to which they have not previously been exposed. Newborns also discriminate among different tastes: They will suck on a sweet solution longer than on sour, salty, or bitter solutions. The sense of touch contains many types of receptors, including ones sensitive to pressure, pain, and temperature. Even the fetus responds to touch, and the newborn is sensitive to pain, for example, by crying following a pin prick. Other aspects of haptic perception such as the detection of shape, texture, hardness, volume, and weight have been found to be functional within the first one to two years of life.

Internal Senses: Orienting and Proprioception

There are other sensory systems besides the five classic senses used for detecting external environmental information. The orienting system is the sensory system that allows a person to detect the position and motion of the body in space. The vestibular organs, consisting of the saccule, the utricle, and the semicircular canals, are primarily responsible for this sense. The orienting system is used to register linear and rotary acceleration. It is also used to sense gravity. Certain types of eye movements show that the vestibular system is functioning. For example, when an individual is rotated, the resulting stimulation of the semicircular canals creates a pattern of eye movements such that the eyes move slowly in the direction opposite of the rotation and then rapidly back. These eye movements are called vestibular nystagmus, and they are easily observable in the newborn infant. With respect to gravity, one interpretation of the frequently observed behavior of infants systematically dropping objects and watching them fall is that it is helping them learn about gravity. Also, the sense of gravity is so strong among 2-year-olds that they search directly under the location where a ball was dropped even in the presence of an opaque tube carrying the ball to a different location. In other words, 2-year-olds do not attend to the local information of the tube; instead they use general information about gravity to attempt to solve these problems.

Proprioception provides information about the position and movement of parts of the body, that is, the status of the muscles, tendons, and joints. Several examples provide evidence for the early emergence of this sensory system. Through ultrasound (high-frequency sound

wave reflections used to see an outline of the fetus), fetuses have been observed opening their mouths in anticipation of the arrival of the arm that allows them to suck on a hand or thumb. Hence, they have some rudimentary knowledge of the location of the hand with respect to the mouth. Imitation of mouth movements among newborns, such as opening the mouth or extending the tongue, seem to indicate a mapping between what is observed in others and how such actions feel to the infant. Recall that newborns cannot see their own mouths, so they cannot visually match what they observe in others with what they observe in themselves. At around 2 months of age, infants are frequently observed watching their hands move (also known as visual hand regard or visual capture), which may help them to develop knowledge of where the limbs are located in space. Finally, infants have looked at monitors that show their moving legs in real time from their own perspective and from another perspective. They discriminate between the two perspectives by looking longer at the unfamiliar perspective. This discrimination requires matching how they feel their legs to be moving with the visual consequences of these movements.

Sensory Coordination

Intersensory Integration

The various sensory systems often do not operate alone. Many objects in the environment stimulate multiple senses, for example, one may see and hear a person, see and touch a toy, smell and taste food, and so forth. Two perspectives on the coordination of information from the sensory systems address whether differentiation or integration occurs during development.

One perspective is that perceptual information is initially undifferentiated, such that a newborn is unable to distinguish which sense is being stimulated. Some speculate that this experience may be similar to the perceptual disorder synesthesia, in which two or more body senses are coupled. In infants, early demonstrations of intersensory coordination support this perspective. For example, 4-month-old infants were simultaneously shown two animations with one soundtrack, which originated from a speaker located in between the two animations. The soundtrack matched the events in one of the animations but not in the other. Infants looked longer at the animation that corresponded to the soundtrack than at the other animation. In other examples, infants were allowed to explore an object with their hands or

their mouth without looking at the object. After the tactile experience, they were shown two objects. One was the object they had just been familiarized with, and the other was a novel object. Their looking times indicate that they discriminated between the two objects. Hence, they can match how the object felt with what it looked like. With experience and development, the child becomes better able to differentiate among the senses and to recognize which sense is being stimulated.

The second perspective is that the senses are initially separate, develop relatively independently, and become connected as they mature. The famous cognitive developmental theorist Jean Piaget argued from this perspective. Specifically, that much of the knowledge and skill underlying intermodal perception can only be gained through experiences that involve looking at, hearing, smelling, tasting, and touching the external world.

A hybrid of these two perspectives is that the senses are initially undifferentiated, then they become increasingly differentiated, and finally they become coordinated. Regardless of the developmental pathway, the ability to integrate information across sensory systems is important in many ways. It enables one to recognize objects across different modalities and to associate sights and sounds. For example, infants learn to associate a particular voice with a particular face by 3 months of age. The ability to localize sound is important both for guiding visual attention and for locating sound-producing objects. Newborns turn their head in the direction of a sound, though this ability briefly declines before reemerging at approximately 4 months. One explanation offered for this pattern is that localization is controlled in a subcortical way at birth, that the cortex takes over as it matures, that it is not developed enough to be very accurate at first, but that it rapidly gets better at localizing sound.

Perceptual-Motor Coordination

Perceptual-motor coordination is the linking of perception and self-initiated movement. James Gibson argued that these two processes are inextricably coupled. For example, in evolutionary terms, people must perceive objects and events in the environment to survive, and survival requires that actions be guided by perceptions. Hence, the purpose of perception may be to guide action, and the purpose of action is to generate additional perceptual information.

One example of perceptual-motor coordination concerns the acquisition of visually guided behaviors, such as reaching. The development of reaching is facilitated by the experience of observing the limbs during active movement. This experience helps one to localize the limb in space (proprioception) and link movement with the visual consequences. It is not clear that the limbs must then remain visible during the targeted action. For example, normally developing infants who are experienced with seeing their limbs move in space can reach in the dark for a glowing object without being able to see their hands.

A second demonstration of perceptual-motor coordination concerns knowledge of the location and distance of sound-producing objects. When placed in the dark, infants reach in the correct direction for a sound, and they are more likely to reach for a sound located within reach than for one located beyond their reach.

Calibration and Recalibration

The linking of various sensory systems and of the perceptual-motor system is not a one-time event. A continuous calibration of the systems maintains accuracy and accommodates changing circumstances, such as the growth of the individual or the transition to more powerful vision correction glasses. Many have argued that vision is the most powerful sense for the process of calibration and that it dominates the other sensory systems. One everyday example of calibration comes from experiencing a moving walkway at the airport. Because the floor is moving as one walks, the effort put into self-movement indicates that one is going faster than that same effort would normally suggest. Hence, if one is then immediately placed in a situation without visual feedback, one may underestimate the distance or time it would take to walk from one location to another. A second demonstration of recalibration occurs when distorted stimuli are presented and people adapt other senses and motor movements to the distortion. For example, distorted prisms placed on the visual system leads to adjustments in pointing and reaching behavior as well as to adjustments in sound localization. These distortions persist for a brief period even after the prisms are removed.

Brain Development and Experience

Role of Experience

One theme that has been mentioned in this entry but not explicitly discussed concerns the roles of

experience on perceptual development. The two extreme positions are maturation and induction. With maturation, experiences play no role in the development of the trait. For example, young infants can differentiate among different odors regardless (apparently) of their previous experiences. With induction, a trait only develops as the result of specific experiences. For example, in a congenitally and completely sightless individual, the visual cortex may not develop because it never receives any visual input. Other concepts, such as maintenance, facilitation, and attunement, take more of a middle position. Maintenance represents the idea of "use it or lose it." The trait develops without any specific experiences but is only retained with experience. For example, infants are sensitive to phonemes in unfamiliar languages, but as they get older they can only detect phonemic differences in the language(s) they hear. Facilitation means that, although specific experiences may hasten the rate of development, others who did not receive those experiences eventually catch up. For example, an enriched visual environment may speed up the development of some visual abilities, but those in a "normal" visual environment would also develop those abilities. The term *attunement* is used to describe an increase in the level of development achieved on the basis of specific experiences. For example, some perceptual-motor coordination may be achieved without ever viewing the limbs, but experience seeing the limbs may allow for a higher level of perceptual-motor coordination.

Brain Development

Developmental changes in the brain may contribute to perceptual development in several ways. Two of these ways include developmental changes in the connections among neurons, and changes in the rate of maturation of different brain structures.

Synaptogenesis is the formation of synapses or gaps between nerve cells (neurons). Within a neuron, the transmission of information is electrical, as the neuron fires in response to certain stimulation. Between neurons, neurotransmitters (chemicals) flow across a synapse allowing neurons to communicate with each other. In many parts of the brain, there is a distinct course of development consisting of the overproduction and pruning of synapses. The proliferation of synapses early in development means that the toddler brain has far more synaptic connections than the adult brain. During childhood, the number of synapses decreases

to adult levels. Experience plays an important role in determining which synapses are maintained and which are pruned. If experiences lead to neurons firing and neurotransmitters being released, then the synapses are maintained. If experiences do not lead to neurons firing and neurotransmitters being released, then the synapses wither.

The process of synaptogenesis is consistent with the hybrid model of the development of intersensory coordination that was presented earlier. During the synaptic proliferation phase, the senses may be undifferentiated, allowing intersensory coordination abilities to be detectable very early in development. Experiences within particular sensory systems then lead to an increased differentiation of the senses and also to a meaningful integration of the senses.

The relative size and level of activity in different areas of the brain changes during development. The cerebral cortex is immature relative to other parts of the brain throughout much of childhood. The cerebral cortex is involved in the processing of perceptual information, as well as many other functions. Thus, despite the emphasis in the literature on the earliest appearance of perceptual capacities, which is the reason this entry has focused on infancy, it is clear that perceptual development is very likely to continue throughout childhood.

A recent theory of visual processing in the brain has a focus on the cerebral cortex. This theory of normal visual processing was largely developed from research on people who had a variety of visual disorders due to brain damage. The theory itself does not directly address development, though many researchers are applying it to developmental issues. David Milner and Melvyn Goodale introduced this theory of the visual brain. Structures in the cerebral cortex relevant to the theory include primary visual cortex, posterior parietal cortex, and inferotemporal cortex. Briefly, two streams of visual information leave the primary visual cortex: A dorsal stream that goes to the posterior parietal cortex, and a ventral stream that goes to the inferotemporal cortex. These visual streams serve different purposes.

The ventral stream is involved in the formation of perceptual and cognitive representations of objects and of their significance. Take, for example, identifying an object. The object's identity is independent of any particular viewpoint of the object. Hence, size and shape constancy enable enduring characteristics of the object to be maintained across different viewing perspectives.

The identity of objects and their spatial arrangements should also be stored in long-term memory to maximize the efficiency of identification.

The dorsal stream is involved in the control of goal-directed actions. Take, for example, reaching out and grasping an object. The location of the object must be specified in an egocentric manner (that is, with respect to the actor) so that the actor can move the hand to the location of the object. The object's size and shape also need to be specified in terms relative to the actor so that the hand and fingers can be adjusted appropriately for grasping. Because actors and the objects they interact with almost continuously change relative location, only a very short memory is required.

Hence, the requirements for a system involved with forming representations of objects are different from the requirements for a system involved with the visual control of action. The first stream involves the world independent of the observer, whereas the second stream involves the actor's actions within the visual world.

The dorsal stream is consistent with Gibson's theory that perception and action are closely linked. The ventral stream provides object recognition capacities that are important for higher order cognitive tasks, and it is consistent with Piaget's theory of the construction of knowledge.

Perceptual Capabilities Beyond Infancy

Perception is used to attend to stimuli and events worthy of detailed processing; to identify what is being perceived; and to locate objects and events for guiding action. The emphasis in this entry was on the early appearance of perceptual capabilities, often in infancy. In contrast, some of the areas mentioned in perceptual abilities among adults, such as the cerebral cortex, are immature in infancy and slowly mature throughout childhood and adolescence. Hence, there is danger in all-or-none thinking about perceptual capabilities. In other words, just because there is some evidence that infants can use certain features or information does not mean that the processes underlying that ability are fully formed. Perceptual development continues well beyond infancy, though there currently is not much emphasis on specifying how that occurs.

Michael E. McCarty

See also Autism Spectrum Disorders; Cognitive Development and School Readiness; Disabilities; Dyslexia; Early Intervention Programs; Motor Development; Observational Learning; Special Education; Working Memory

Further Readings

Bukatko, D., & Daehler, M. W. (1995). *Child development: A thematic approach* (2nd ed.). Boston: Houghton Mifflin.

Gibson, J. J. (1979). *The ecological approach to visual perception.* Boston: Houghton Mifflin.

Greenough, W. T., Black, J. E., & Wallace, C. S. (1987). Experience and brain development. *Child Development, 58,* 539–559.

Haith, M. M. (1980). *Rules that infants look by.* Hillsdale, NJ: Lawrence Erlbaum.

Milner, A. D., & Goodale, M. A. (1995). *The visual brain in action.* New York: Oxford University Press.

Piaget, J. (1971). *The construction of reality in the child.* New York: Ballantine.

Schiffman, H. R. (2000). *Sensation and perception: An integrated approach* (5th ed.). New York: Wiley.

Schone, H. (1984). *Spatial orientation: The spatial control of behavior in animals and man.* Princeton, NJ: Princeton University Press.

Siegler, R. S., & Alibali, M. W. (2004). *Children's thinking* (4th ed.). Upper Saddle River, NJ: Prentice Hall.

PERSONALITY TESTS

It has long been observed that individuals differ one from another on many psychological dimensions. An area of intense interest among psychologists is the measurement of individual differences in personality. Personality is commonly defined as the constellation of traits, or typical and relatively stable patterns of responding to the environment, which are unique to various individuals. An important focus of educational psychology is the assessment of these traits and other related psychological attributes such as interests, preferences, and attitudes. This entry provides a brief overview of the history of personality testing, examines important ethical and psychometric issues related to the use of personality tests in clinical and educational settings, and describes the most commonly administered personality tests.

History of Personality Testing

Documented use of some form of testing in an effort to place and classify individuals dates back to at least

2200 B.C., when the ancient Chinese used such tests to determine the placement of individuals in various civil service positions. Plato and Aristotle noted that humans differ in personality traits, interests, and intellectual abilities and that these differences can be assessed in various ways. However, it was not until the late 19th century that interest in the scientific measurement of human differences began. In 1869, the British psychometrician and statistician Sir Francis Galton (cousin of Charles Darwin) published the influential *Classification of Men According to Their Natural Gifts.* This text helped lay the groundwork for the field of psychological measurement by focusing attention on precision and accuracy in measurement and by advancing several basic statistical techniques, such as correlation and regression. At approximately the same time, the scientific study of human mental processes was being advanced by German psychologists such as Wilhelm Wundt, who is generally acknowledged as the founder of experimental psychology. In this new field, scientific observation began to replace subjective impression and introspection as the primary means by which to study mental processes. James McKeen Cattell, a student of Wundt and an acquaintance of Galton, brought the German emphasis on measurement precision and Galton's work on psychological testing to American academe. Another important advance was the 1904 publication of the first textbook on educational measurement, *Introduction to the Theory of Mental and Social Measurement*, by E. L. Thorndike. The major accomplishment of this early era, however, was the 1905 publication of the first standardized psychological test, the Binet–Simon Intelligence Scale, by French psychologists Alfred Binet and Théodore Simon. This test ushered in the modern era of standardized testing, and no other single test has had a greater influence on the field of psychological and educational measurement than the Binet–Simon IQ test. In fact, its revision by Stanford psychologist Lewis Terman in 1916 resulted in the still used and widely adopted Stanford–Binet Intelligence Scale, currently in its fifth edition.

The assessment of personality traits and related psychological constructs parallels the development of the major theories of personality and important advances in psychometric and statistical techniques. For example, Freud's psychoanalytic theory, with its emphasis on unconscious needs and drives, inspired the development of two of the most well-known projective tests: the Rorschach Ink Blot Test (1921) and the Thematic Apperception Test (1938). Another example is the more contemporary Myers–Briggs Type Indicator (1985), based on personality types proposed by Carl Jung, the Swiss neo-Freudian. Trait theorists, such as Raymond Cattell, used sophisticated statistical methods such as factor analysis to reduce lists of hundreds of trait descriptors to the 16 most basic ones, which are termed *factors.* The product of his work, the Sixteen Personality Factor Questionnaire (16 PF), is now in its fifth revision and is considered one of the most psychometrically sound and useful measures of personality traits in use today. Perhaps the most famous of the empirically (or statistically) derived personality tests is the Minnesota Multiphasic Personality Inventory (MMPI), which was developed by S. Hathaway and J. McKinley in the early 1940s. Hathaway and McKinley used a psychometric technique called *empirical criterion keying*, in which scores are computed based on their differentiation of individuals according to some external criterion. In the case of the MMPI, responses on each of the scales distinguished various groups of psychiatric patients from control groups of nonpsychiatric individuals. Another example of a widely used empirical criterion-keyed test is the Strong Interest Inventory (SII), originally developed in the 1920s by E. K. Strong. During test development of the SII, Strong found that individuals who worked in different occupations tended to respond in distinctive ways to various questions relating to their everyday likes and dislikes; subsequently, responses to these questions on the inventory were empirically keyed for different occupations.

Personality testing today is used in a variety of settings for a multitude of purposes. In clinical settings, it is used primarily to assess for normal and abnormal personality functioning and to monitor change during the course of psychological treatment. In educational and business settings, personality and interest inventories are used to evaluate career interests and preferences, learning styles, and academic areas of interest. In legal settings, personality tests may be used by forensic practitioners as part of competency and custody evaluations, assessment of dangerousness and malingering, posttreatment/release decisions, and other such related uses. In research settings, personality and related measures are used extensively in both basic and applied research in the psychological sciences. In fact, a review of the psychological research literature over the past 50 years reveals that the MMPI alone has been used in more than 6,000 published studies. Hundreds of standardized personality tests have been developed,

and they are second in number only to achievement tests.

Ethical Use of Personality Tests

Because of the sensitive nature of the items (e.g., "I have tried to commit suicide") and test results obtained on personality measures, great care is taken to ensure that they are used appropriately and ethically. The overriding goal is to encourage respect for the individual taking the test and respect for the tests themselves. More specifically, the professional standards are intended to keep unqualified individuals from using and misusing such tests. Professional organizations, such as the American Psychological Association and the American Educational Research Association, have guidelines for the ethical use of psychological and educational tests. Examples of standards include

- proper training for those who use such tests;
- treating test results and other personal information as "privileged communication" and keeping them confidential (i.e., in educational settings, only school personnel who have a "legitimate educational interest," e.g., a school psychologist or guidance counselor, may have access to test records);
- obtaining informed consent of the examinee (or if a minor, parental consent);
- maintaining test security (to prevent the test from being disseminated to the general public); and
- using the "least stigmatizing label" for an individual when interpreting results.

Although this list is not inclusive, it illustrates the professional and ethical obligations required of those who administer and interpret psychological tests.

Test-Reliability and Validity

For a test to be a useful measure of a construct, it must, at minimum, measure that construct consistently. The consistency, or accuracy of a measure, is termed *test reliability*. As an example, imagine a bathroom scale that produces substantially different readings when stepped on several times by the same person wearing the identical clothing within 1 minute's time. The fluctuation in readings reflects inaccuracy (or, more literally, inconsistency) in the scale. Under these circumstances, it would be difficult to know the "true" scale reading of an individual's

weight. Psychometricians would say that this particular scale contains a substantial amount of *measurement error*, which, in turn, contributes to its lack of reliability. In contrast, precision measures in the physical sciences, such as thermometers that consistently and accurately read to the millionth of a degree Celsius, are considered highly reliable measures primarily because they have a minimal amount of measurement error. However, in the behavioral sciences, such precision is not possible for a variety of reasons, perhaps the most basic being the inherent variability in human behavior. Similarly, any test condition (e.g., slight differences in the testing environment, or fluctuations in mood states of examinees) that is not relevant to the purposes of the test contributes to measurement error. It is simply not possible in the psychological sciences to construct perfectly reliable tests, so the goal is to construct tests in such a way that measurement error is minimized and test reliability maximized.

One of the most common methods of assessing reliability is to compare the scores of groups of individuals on a test taken at one time with their scores on the same test, or an alternate form of the test, taken at another time. This is referred to as *test–retest reliability*. If the test is reliable, the association, computed as a correlation coefficient, between the two sets of scores should be fairly high. A correlation coefficient is a number that ranges from -1.0 to $+1.0$, with higher numbers reflecting a greater degree of association between two sets of scores. It is desirable in psychological and educational measurement for test–retest correlation coefficients to be higher than .85. Test–retest reliability coefficients and the time interval between test administrations should always be reported in the test manual.

Another method of assessing reliability, particularly when an alternative form of a test is unavailable or is too expensive to construct, is *internal consistency reliability*. In this form of reliability, the consistency of individual test questions (or *items*) is examined from a single administration of the test. For example, in the *split-half method*, a group of test scores from one test is split in half, with the scores on the first half correlated with the scores on the second half or with the scores on the odd-numbered items correlated with the even-numbered items. The higher the correlation is between the two halves of the test, the better is the internal consistency of the test. In *coefficient alpha reliability*, the consistency of all, not just half, of the items on a test are intercorrelated at

one point in time. Coefficient alpha produces a value that ranges from 0 to 1.0, with higher values reflecting greater consistency. This *inter-item consistency* reflects measurement error from two sources: (1) the content of the items being sampled (as in split-half reliability), and (2) the extent to which all of the items measure the same thing, or what is often termed test or scale *homogeneity.* All other things being equal, the more homogenous the behavior being sampled is, the higher the inter-item correlations will be and hence the better the internal consistency of the measure will be. For example, if a personality test contains items that sample very different behavioral traits, such as introversion, aggressiveness, friendliness, or conscientiousness, that test will likely not be a homogenous measure of personality and will produce a low coefficient alpha estimate of reliability.

The concept of *test validity* refers to what the test actually measures and how adequately it does so. In other words, validity examines to what extent the test accurately measures what it purports to measure. For example, suppose an individual takes an introductory psychology course and studies relevant material for the final exam. However, when the individual takes the exam, he or she is asked questions about economics, history, and calculus. Quite obviously that test would not be a valid measure of introductory psychology content and should not be used for that purpose. It is also important to note that tests cannot be valid if they lack reliability; thus, reliability is said to be a "necessary but not sufficient condition" for test validity. Assessments of test validity are complex and always require taking into consideration the particular purpose of the test and the degree to which it fulfills that purpose. Two kinds of validity procedures, both of which are relevant to personality tests, are *criterion-related validity* and *construct validity.*

Criterion-related validity entails evaluating performance on a particular test against a separate criterion, such as teacher or supervisor ratings, clinical diagnoses, or parent reports. The criterion should be one that is logically related to the construct being assessed by the to-be-validated test. For example, scores on the various scales of the MMPI should be more highly related to psychiatric diagnoses than to intelligence test scores. *Concurrent validity* is when the criterion against which the test scores are validated is measured at approximately the same time. An example would be MMPI scores correlated with currently available psychiatric diagnoses. When the criterion is not

available until a certain amount of time has passed, then *predictive validity* is involved. The question of interest with predictive validity is to what extent do the test scores predict performance on a related measure sometime in the future? For example, a test of career interest could be given to a group of individuals at one point, and then those same individuals could be evaluated later with respect to the careers they actually pursued. With both concurrent and predictive validity, the relation between the criterion and the test scores are expressed as a validity coefficient, with higher coefficients suggesting greater validity. Personality tests are often validated concurrently and predictively against such criteria as clinician and teacher ratings of personality traits, psychiatric diagnoses, various indices of job performance, and participation in certain activities (e.g., social or religious activities).

Construct validity is perhaps the most important type of validity for psychological tests, particularly measures of personality. This type of validity concerns the extent to which a psychological test measures a theoretical construct, such as anxiety, ego development, introversion/extraversion, or intelligence. In fact, all other validation procedures fall under the broad umbrella of construct validity. Researchers examine various types of data that may bear on construct validity—for example, correlations with other theoretically related and unrelated measures; differences in trait scores among groups of individuals of different ages or among previously identified groups of people (e.g., those diagnosed with an anxiety disorder compared with those diagnosed with a psychotic disorder); or the pattern of correlations among multitrait (many traits measured by same method) and multimethod (the same traits measured by different methods, such as self-report vs. teacher rating) analyses of the construct. In sum, this type of validity takes all that is currently known about a test and its related construct and evaluates these different sources of information for evidence of its theoretical validity.

Self-Report Personality Tests

The majority of personality tests and interest inventories are paper-and-pencil, self-report questionnaires. Self-report means that test-takers are literally reporting on themselves; that is, they are asked various types of questions about themselves. This technique capitalizes on the fact that individuals generally do

know themselves best and are normally truthful in their responses. However, it is not without its limitations, particularly when individuals are motivated (consciously or not) to respond in a less than truthful manner. Several such *response sets* have been identified, and on most major personality tests, special validity scales are used to detect these response tendencies. For example, many tests have scales that detect or correct for the tendency to present oneself in an overly positive light (so-called social desirability). Other tests have validity scales that detect malingering, or "faking bad," as might occur when there may be a secondary gain associated with the assessment outcome (e.g., needing to appear mentally ill in order to qualify for disability). The various validity scales cannot completely screen out all invalid test protocols, but they do a reasonably good job of detecting the majority of them.

Minnesota Multiphasic Personality Inventory–2

The MMPI–2, originally developed in 1943 via empirical criterion keying, is a 567-item true–false inventory consisting of 10 clinical scales and 3 validity scales. There are also over 20 supplementary scales that measure additional response sets and areas of clinical concern, such as addiction potential and marital distress. The clinical scales include Hypochondriasis, Depression, Hysteria, Psychopathic Deviate, Masculinity–Femininity, Paranoia, Psychasthenia, Schizophrenia, Mania, and Social Introversion. The validity scales are the Lie Score, Infrequency Score, and Correction Score. The MMPI–2 was normed on a sample of 2,600 male and females, 18 to 80 years of age, from various ethnic groups and diverse geographic regions. It uses a standard T score (mean $= 50$, standard deviation $= 10$), with clinical cutoffs for each scale set at 65T. That is, a score higher than 65T would indicate an area of clinical concern. As an alternative to the laborious hand-scoring method, computerized scoring software is provided by the publisher. The resulting scoring profile is then interpreted in terms of the pattern of elevations across all of the scales, although there are more complex interpretive procedures available from different test researchers. The interpretation of an individual's profile only provides possible hypotheses about how that person functions across many domains in different settings. (It is important to note that psychological tests should be used only as adjunct sources of information about an individual and never in isolation.) Among the many uses of the MMPI–2 are to assess for symptoms of psychopathology and personality adjustment, to aid in screening personnel for high-risk public safety jobs, and to support treatment and management decisions in forensic settings. Without question, the MMPI has been the most extensively used personality test in both clinical and research settings, worldwide.

Sixteen Personality Factor Questionnaire

The 16 PF, originally developed in 1949 by R. B. Cattell, and now in its fifth revision, uses the *factorial approach* to test construction. Factor analysis is a data-reduction technique that analyzes sources of variance among test scores such that only those dimensions (or factors) that account for the largest portion of the common variance are identified. Using this approach, Cattell identified the 16 basic or most primary personality traits (e.g., Warmth, Social Boldness, Perfectionism, Emotional Stability, Sensitivity). The 16 PF also provides five global factor scores (Extraversion, Anxiety, Tough-Mindedness, Independence, Self-Control) and three validity indices (Impression Management, Infrequency, and Acquiescence). The current edition, consisting of 185 multiple-choice questions and appropriate for individuals 16 years and older, is used by psychologists for assistance with clinical diagnosis, treatment planning, and prognosis; vocational and career counselors for assessing occupational interests and preferences; marriage and family therapists in evaluating marital compatibility and satisfaction; and school psychologists and counselors to help identify students with possible social, emotional, or academic problems. It was normed with a stratified random sample based on the 2000 U.S. Census and consisted of more than 10,000 individuals, age 16 and older. In general, the various scales of the 16 PF have demonstrated very good estimates of reliability and moderate to good indicators of validity. Computerized scoring and score reports are available from the publisher.

Strong Interest Inventory

The SII was first published in 1927 by E. K. Strong, a Stanford University psychologist interested in understanding the characteristics of people in different kinds of occupations. Using an approach that was unique in

test construction at the time, he first asked people about their everyday likes and dislikes with respect to different kinds of activities. He then used empirical keying to identify the pattern of responses associated with a particular occupation. His use of empirical keying to derive the items and scales predates even the well-known empirically keyed MMPI. Today, the SII is one of the most popular interest inventories and is used by guidance and career counselors in secondary and higher education settings and in business settings to improve workplace productivity and employee retention. The SII has recently been re-normed and revised to include several new interest and occupational scales. The test now contains four major scales (General Occupational Themes, Basic Interest Scales, Occupational Scales, and Personal Style Scales), each with several subscales. The scales use a standard T score, with a mean of 50 and standard deviation of 10. The SII can be scored only by computer. The basic report displays the examinee's scores on all of the scales and subscales, and an interpretive report provides information about specific careers and occupational interests. Overall, the SII has an excellent psychometric track record with solid reliability and validity.

Projective Personality Tests

Among the educated lay population, the most widely familiar personality tests are the *projective tests.* Through the media, such as popular television sitcoms like *Fraser*, most people have likely heard of the Inkblot Test. The *projective hypothesis*, a term first proposed by psychoanalyst Lawrence Frank in 1939 and based originally on Sigmund Freud's early work on projection, assumes that people interpret ambiguous stimuli in terms of their unconscious inner drives, conflicts, and concerns. That is, they "project" onto vague and ambiguous unstructured stimuli what they really feel, think, and worry about. As such, these tests are considered a type of *disguised* assessment in which the test-taker is generally unaware of the purpose of the test or the possible psychological interpretations derived from it. Proponents of this technique argue that because they are a type of disguised test, they are difficult to fake. However, others have argued that precisely because the stimuli are so unstructured, responses to them can reflect an array of processes: conscious and unconscious, significant and meaningless, real and imaginary. Thus, interpretation of the responses is often subjective and unreliable,

a factor that undoubtedly contributes to the generally low validity coefficients of most projective instruments. Given space limitations, only the two most commonly given projective tests, the Rorschach Inkblot Test and the Thematic Apperception Test, are discussed here.

Rorschach Inkblot Test

The Rorschach, originally published in 1921 by the Swiss psychiatrist Hermann Rorschach, and currently the most popular projective measure among clinical psychologists and psychiatrists, consists of a series of ten inkblots (five black and white, two red and gray, and three multicolored). Individuals are simply asked to tell the examiner what they might see in each blot. After the initial responses to all the cards, the examiner may readminster the cards to probe the test-taker for specific responses to various features of the cards. Several scoring methods have been proposed, the most recent of which is John Exner's comprehensive system. This very complex approach (simplified here) attempts to quantify the scoring of the responses as much as possible by evaluating every response to each blot on several dimensions:

1. Location (e.g., use of the whole blot or only a part of it)

2. Determinant (e.g., use of color, form, or shading in the response)

3. Content (e.g., references to anatomy, nature, or fire)

4. Popularity (whether the response is common or original)

The number and ratio of responses in each category are then interpreted according to Exner's guidelines. For example, a large number of color responses indicates emotional lability and impulsiveness, whereas many detailed responses might indicate an overly controlled or compulsive personality. Although Exner's system represents a substantial improvement over previous scoring systems, there is still a considerable amount of debate in the scientific community about the adequacy of the Rorschach's psychometric properties. Specifically, while there is some support for the use of a small number of the specific indices found on the Rorschach (e.g., Schizophrenia Index), the majority of studies indicate that the Rorschach

fails to meet many of the commonly accepted standards of reliability and validity. Nonetheless, it continues to be widely used by clinicians.

Thematic Apperception Test

The Thematic Apperception Test is the second most commonly administered projective test. It was developed by Henry Murray and his staff at the Harvard Psychological Clinic in 1938 and comprises 19 black and white picture cards that depict scenes of people in different ambiguous situations. (One card is blank.) Individuals are asked to tell a story about what is happening in the various scenes (with the blank card they are asked to imagine a scene and tell about it), what the characters are thinking and feeling, what led up to the story, and what the outcome might be. For example, one card depicts a young boy who appears to be contemplating a violin that sits on a table in front of him. An individual might respond with a story about a little boy who feels conflicted about learning the violin and is torn between pleasing his parents and doing what he really wants to do, for example, going outside to play with his friends. Ultimately, he decides to pick up the violin and commence practicing. Clinical interpretation of the stories requires ascertaining the main motives, needs, conflicts, and emotions projected onto the characters in the pictures by the examinee. The "hero" or "heroine" of the stories is assumed to be the examinee, with conflicts conveyed in the stories reflecting putative stressors in the individual's life. Although the Thematic Apperception Test is widely used in clinical settings, it has many of the same psychometric limitations as the Rorschach, although in general it fares somewhat better in terms of validity.

Tracie L. Blumentritt

See also Generalizability Theory; Measurement

Further Readings

Aiken, L. R. (2000). *Psychological testing and assessment* (10th ed.). Boston: Allyn & Bacon.

Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice Hall.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*, 281–302.

Exner, J. E., Jr. (1993). *The Rorschach: A comprehensive system: Vol. 1. Basic foundations* (3rd ed.). New York: Wiley.

Hopkins, K. D. (1998). *Educational and psychological measurement and evaluation* (8th ed.). Boston: Allyn & Bacon.

Lillienfeld, S. O., Wood, J. M., & Garb, H. N. (2000). The scientific status of projective techniques. *Psychological Science in the Public Interest*, *1*, 27–66.

Murphy, L. L., Conoley, J. C., Impara, J. C. (Eds.). (1994). *Tests in print IV* (2 vols.). Lincoln, NE: Buros Institute of Mental Measurements.

Spies, R. A., & Plake, B. S. (Eds.). (2005). *The sixteenth mental measurements yearbook.* Lincoln, NE: Buros Institute of Mental Measurements.

Personalized System of Instruction

The Personalized System of Instruction (PSI) is a teaching method designed by Fred S. Keller and J. Gilmour Sherman from Columbia University. It focuses on mastery of knowledge or skills, as reflected in examination performance, and is implemented typically at the college level, although reports of high school and upper elementary implementation exist. Students must have successful reading, writing, and study skills to profit from PSI. Inspection of journals and magazines devoted to teaching methods in various disciplines indicates that PSI courses have been offered in more than 50 disciplines, including psychology, engineering, oceanography, English, sociology, biology, physics, history, and philosophy. PSI is derived from principles of behavior analysis, and has five distinguishing features. This entry describes those features and the outcomes of PSI research.

Stress on the Written Word

Communication between the teacher and the learner is by written word, usually the standard textbook, computer-generated documents or Web sites developed by the instructor. Many distance learning courses are designed in a similar fashion; PSI is an ideal format for offering a web-based course. The PSI instructor prepares detailed study guides by listing all of the learning objectives for a reading assignment. For example,

On completion of this unit on the basic principles of PSI, be able to

1. describe five distinguishing features of PSI,

2. analyze written descriptions of PSI courses, and suggest improvements based upon the five distinguishing features,

3. give an original example in which you describe a new PSI course; your choice of subjects.

   (and so on)

Many PSI study guides contain study questions that follow along with the text, prompting active responses, instead of learning objectives.

Unit-Perfection Requirement for Advancement

PSI instructors divide the content of their courses into one to two subdivisions or "units" for each week of the course. Once instructors outline a unit sequence, they develop three to four parallel versions of quizzes for each unit. Unit quizzes usually contain multiple-choice and short-answer essay questions and typically take 10 to 20 minutes to complete. Students take tests on each unit as many times as necessary until they reach "mastery," which is generally defined as 90% correct. Students are not penalized for errors. As Sherman often said, we judge our art masterpieces not by the number of preliminary sketches discarded along the way, but by the final product produced. Once students achieve a mastery score on a unit quiz, they proceed to the next unit.

In traditional classrooms where mastery is not the goal, time to learn objectives or a module or topic is held constant and, as a result, quality of learning varies. Both time and quality cannot be held constant. At the end of a time period, a teacher may test students' competencies, or projects may serve as evidence of progress. Progress will vary from the most to least "capable" learners. In contrast, in PSI, quality of learning is held constant and time is allowed to vary. As with other mastery learning approaches of the 1960s and 1970s, PSI advocates argue that the aspect of teaching that should vary is "time to completion" of a course of study, not "quality of learning." The mastery philosophy of learning recognizes that individual differences exist among students and that these differences are reflected in the learning process. Some learners may need extra test opportunities for a variety of reasons, including learning challenges, family circumstances, and other demands on their lives and

times. But these differences are more adequately described by differences in rate of learning, not by the final level of learning achieved.

Student Pacing

Student pacing allows each individual to proceed through the unit sequence in a PSI course as slowly or as rapidly as possible, achieving virtual mastery of the course material in either case. Students are neither left behind nor painfully bored. If circumstances arise that make effective study impossible, or if students have real difficulty learning the material, they can delay taking a unit quiz, and progress more rapidly at a later point in the course.

Use of Proctors

PSI instructors enlist the services of students who have already mastered course content to serve as proctors. The proctor (a) immediately scores and evaluates students' performance on successive quizzes, (b) indicates any relevant portions of material that a student may need to restudy before retaking a quiz, (c) suggests ways of improving study behaviors and examination skills, (d) prompts consistent progress through the course, and (e) "personalizes" a PSI course. Indeed, the proctor staff in many ways determines the success or failure of a PSI course. Most PSI courses recruit proctors to create a 1:10 proctor/student ratio.

External Proctors

The proctor is typically an undergraduate or graduate student who has been chosen for his or her mastery of the course content and positive approach to the field of study. Eighty percent of all PSI courses use external proctors. Students who complete a PSI course at a high rate are often recruited for proctoring during a subsequent term. The instructor and current team rate potential student proctors on characteristics such as "sociability," "dependability," "knowledge of materials," and "overall ability to be a manager." They also review the students' quiz folders. Those who are rated highest may then be invited to be proctors. External proctors typically receive two to three hours of "independent study" course credit. The proctor's grade may be based on a number of factors, including class attendance and

performance on an exam over course content. Some departments establish a "Teaching Assistant" course number for proctors and may require their participation in a seminar.

Internal Proctors

In some PSI courses, students concurrently enrolled in the course serve as internal proctors. Students who have successfully mastered the greatest number of units may volunteer their services as proctors for that day. In a class of 100 students, the 10 students who serve other students best are chosen. These proctors then score all units up to the last unit that they have mastered. In this system virtually everyone has a chance to proctor by gaining the lead in progress through the units in the course. Slower students get their chance when the first wave of students completes the course. The instructor must maintain more direct contact and closer involvement with internal proctors than external proctors, due to their relative lack of sophistication in handling student difficulties with course content.

Many PSI courses have a two-tiered proctoring system, in which external proctors are used in a 1:20 proctor/student ratio; they supervise additional, daily internal proctors, thereby reducing the ratio to the ideal of 1:10.

Lectures and Demonstrations as Motivational Devices

If lectures are interesting and informative, they can be important and highly reinforcing events. Since the PSI instructor uses study materials, self-paced evaluation, and mastery to present essential information, lectures can be based on material that is not absolutely essential to the actual course content but is interesting and informative in its own right. Such lectures can be a refreshing change from the usual course lecture that presents essential but often routine or boring course content. The lecture can become a pleasurable experience for both student and instructor. To promote unit progress, the lecturer can require that students must reach a certain level or pass a predetermined number of units before they are allowed to attend the lecture. The lecture in a PSI course is not eliminated, as in many other forms of individualized instruction, but serves a different and highly useful function.

Three Levels of Evaluation

Research on the effectiveness of PSI has examined student grade distributions, studies comparing PSI with conventional teaching methods, and component analyses of PSI.

Student Grades

Numerous descriptive reports of PSI in college classrooms have described large differences in the final grade distributions of PSI and more traditionally taught classes. Students in PSI courses receive a much greater proportion of A and B grades than do students in conventionally taught sections of the same course. Some reports have also indicated that, in comparison to previous terms, the same professor assigns many more As and Bs when using PSI to teach the same course. The effectiveness of the explicit contingencies given to students in a PSI course (what to learn, how to show it on a test) appears to produce a ceiling on grades awarded, especially when grades are based on tests. Some reports note a higher incidence of "Incomplete" grades, and "Withdrawals" in their PSI courses, but researchers have not reported systematic differences.

Comparison Studies

Several reviews have indicated that average student performance in PSI courses ranges 8% to 15% better than student performance in traditionally taught courses. For example, one review of 39 studies comparing PSI to lecture–discussion methods reported that 34 studies (87%) showed that PSI students performed statistically higher on exams. Four studies favored PSI but were not statistically significant; only one favored lecture/discussion, but the results were not statistically significant. A more recent meta-analysis produced similar results. As a point of comparison, 51% of 88 studies on college teaching favor the lecture method, whereas 49% favor a discussion method. No one has reported that any other college teaching method produces better exam performance than PSI.

Nine retention studies have compared PSI to conventional instruction; 100% showed that PSI produced statistically higher retention on exams. Four studies have shown statistically higher grades in subsequent courses after taking an introductory PSI course; one study showed higher grades after a PSI course, but the results were not significant. No studies have

shown that a conventionally taught course produces statistically higher retention or transfer when compared to a PSI course. Nine studies have compared student ratings in PSI and conventional courses. Seven statistically favored PSI ratings, one favored PSI but was not significant, and one favored a lecture/ discussion method but was not significant.

Component Analysis

The mastery criterion may represent the most powerful contingency in PSI courses. Studies have shown that student performance levels will typically match whatever mastery criterion an instructor sets—70%, 80%, or 90%—probably because grades are powerful incentives for most students. Mastery criteria are not as effective in the absence of small units, frequent quizzing, and immediate feedback. Student performance varies inversely with increased unit size, although advantages of smaller unit size are clear only when the differences in unit sizes compared are large.

Restricting student pacing, either with a minimum rate line for unit completion, or with deadlines for early or all units, has shown no effects on final exam performance, although decreases in "Incomplete" and "Withdrawal" grades have been noted. Students highly favor flexible pacing conditions, and choose to perform under such conditions when given the opportunity. Student pacing may also teach self-scheduling skills, although no research has investigated this possibility. The vast amount of research on the pacing component of PSI appears to be the result of problems generated by the incompatibility between self-pacing and traditional university, college, and other institutional policies.

The provision of study questions has a powerful effect on student generalization and recall performance. Monitoring student answers to study questions before they take quizzes improves student pacing and first-trial quiz performance. Grading study-question answers instead of quizzes produces lower final exam performance, probably because generalization and maintenance of performance is better when students must respond to new quiz exercises. Both oral and written quizzes appear to result in high-quality student performance, but the comparative results are not conclusive.

PSI courses are usually more effective with a review procedure. The use of either review items on unit-mastery quizzes or separate review tests has demonstrated positive effects on student performance and retention. The relative gains of elaborate review unit programming have not been investigated. When a mastery criterion in PSI courses is shifted from 90% correct to a time-based measure—for example, six math word problems solved correctly in 3 minutes— retention dramatically improves. Review items or units may be superfluous when a mastery criterion is both time- and accuracy-based.

Immediate feedback on quiz performance from proctors also has powerful effects on student exam performance, although it is difficult to separate the effects of immediate feedback from the use of proctors. Training proctors in administrative, social, accuracy, and prompting behaviors increase those behaviors in the PSI classroom, and the positive effects of proctor prompting and accuracy behaviors on student performance have also been demonstrated. In the absence of proctor training, internal and external proctoring appear to be equally effective. Students prefer feedback from proctors to other feedback arrangements; they also appear to learn more when they engage in proctoring. One report on the use of internal proctors showed that proctor evaluations of their proctoring opportunities were consistently positive. Internal proctors showed significant shifts toward choosing the particular discipline as their major area of study, reported greater interest in the course as compared to their other courses, indicated a high likelihood of returning in a later term to be an external proctor, and stated that they will probably use these same procedures to teach their classes, if they become teachers. Comparisons of trained versus untrained; external, internal, and self-proctoring; on student performance and preference need to be conducted to determine the most optimal proctoring procedures.

Kent Johnson

See also Applied Behavior Analysis; Distance Learning; Effective Teaching, Characteristics of; Operant Conditioning; Teaching Strategies

Further Readings

Johnson, K. R., & Ruskin, R. S. (1977). *Behavioral instruction: An evaluative review.* Washington, DC: American Psychological Association.

Keller, F. S. (1968). "Goodbye, Teacher . . ." *Journal of Applied Behavior Analysis, 1*, 79–89.

Kulik, J., Kulik, C.-L. C., & Cohen, P. A. (1979). A meta-analysis of outcome studies of Keller's personalized system of instruction. *American Psychologist, 34*, 307–318.

Sherman, J. G., Ruskin, R. S., & Semb, G. B. (Eds.). (1982). *The Personalized System of Instruction: 48 seminal papers.* Lawrence, KS: TRI.

Web Sites

Cambridge Center for Behavior Studies: http://www.behavior.org

PHONICS

Phonics is the correspondence between letters and spoken sounds, or letter–sound correspondences. Today virtually all approaches to early reading instruction in the United States include instruction in phonics. The various approaches differ in the units of sound taught and how they are taught. Some teach letter–phoneme correspondences and others teach letter–onset and letter–rime correspondences. Some teach phonics out of context and others teach it in context. However, all approaches share the same goal: to help beginning readers become independent readers.

Spoken Sounds

In English, the sounds employed in phonics instruction can be phonemes or they can be onsets and rimes.

Phonemes

Traditionally, the unit of sound used in phonics instruction has been the phoneme. The phoneme is the smallest unit of speech that makes a difference in the meaning of a word in a given language. For example, in English, if we drop the first /s/ in the spoken word *smiles*, we have *miles*. All languages have phonemes, but the range of sound in any given phoneme differs from language to language. For example, /b/ and /v/ are distinct phonemes in English but variations of the same phoneme in Spanish.

Children who have not yet learned to read have trouble analyzing spoken words into phonemes. Adults who have learned to read can analyze spoken words into phonemes when the phonemes correspond with the spelling (as in *cat*) but have trouble when they do not (as in *box*). (The three letters of box represent four phonemes /b/, /o/, /k/, and /s/.)

Onsets and Rimes

In the 1970s, linguists discovered that spoken syllables in English consist of two natural parts: (1) any consonants before the vowel and (2) the vowel and any consonants that come after it. They named the first part the *onset*, and the second part the *rime*. In the spoken word *smiles*, for example, /sm/ is an onset and /īlz/ is a rime. Syllables may or may not have onsets but all syllables have rimes. For example, the word *smiling* has two syllables and, hence, two rimes but only one onset.

Although many languages have onsets and rimes, not all languages do. The psychological unit of the Spanish syllable is the syllable, not onsets and rimes.

Onsets and rimes may consist of one phoneme (as the /g/ and the /o/ in the word *go*) or more than one phoneme (as the /s/ and /m/ and the /ī/, /l/, and /z/ in the word *smiles*). In words composed entirely of one-phoneme onsets and rimes (as *go*), there are as many units of sound at the onset–rime level as there are phonemes. In all other words there are fewer units at the onset–rime level than phonemes. The spoken word *smiles*, for example, has two units at the onset–rime level (/sm/ and /īlz/) but five units at the phonemic level (/s/, /m/, /ī/, /l/, /z/). Most words have more than one phoneme in the onset or the rime in at least one syllable.

Onsets and rimes are so intuitive to native speakers of English that linguists call them the psychological units of the syllable. Long before linguists discovered onsets and rimes, poets and educators used them but called them by different names. Poets speak of alliteration and rhymes, and educators speak of word families. When there is more than one phoneme in an onset or a rime, English-speaking children who have not yet learned to read are able to analyze spoken words into their constituent onsets and rimes when they cannot analyze them into their constituent phonemes. For example, they can analyze the spoken word *smiles* into /sm/ and /īlz/ but not into /s/, /m/, /ī/, /l/, and /z/.

Letter–Sound Correspondences

Letter–Phoneme Correspondences

In English, the relationship between letters and phonemes is complex. Most letters, either alone or coupled

with other letters, can represent more than one phoneme. For example, the letter *s* can represent /s/ (as in the beginning of the print word *smiles*), /z/ (as in the end of the print word *smiles*), and /sh/ (as in the print word *sugar*). Sometimes a combination of two or more letters, known as digraphs, can represent a phoneme (as the *ph* in *phonics*, the *th* in *the*, the *ght* in *light*, the *oa* in *boat*, and the *oo* in *book*). Sometimes a letter *e* at the end of a print word can influence the pronunciation of a noncontiguous *a, e, i, o,* or *u* that occurs two letters before it (as the *e* in the print word *smile*). In one case a letter represents two phonemes: the letter *x* after a vowel represents /k/ and /s/ (as in the word *box*). Sometimes letters are pronounced in some words but are "silent" in others. For example, the letter *m* is pronounced in the print word *smile* but not in the print word *climb*. The letter *l* is pronounced in the print word *smile* but not in the print word *walk*.

Sometimes the pronunciation of print words is influenced by factors other than letter–phoneme correspondences. In words known as homographs, the pronunciation depends on the word's grammatical function (as the word *read* in "I didn't read this but I read that"). In some words the spelling is morphological rather than phonetic. That is, the spelling represents meaning, not pronunciation (as the *ed* in the print word *dropped*). The pronunciation of print words is also influenced by the reader's dialect. For example, depending on one's dialect, the word *again* can rhyme with *ten* or with *rain*.

Conversely, when converting speech to print, often a phoneme can be represented by more than one letter or set of letters. The phoneme /k/, for example, can be represented by the letter *k* (as in *kangaroo*) and by the letter *c* (as in *cat*). The phoneme $\bar{oo}$ can be represented by the letters *o, ew, oo, ue, iew,* and *ough* (as in *to, blew, too, blue, view,* and *through*).

Phonics generalizations based on letter–phoneme correspondences are not reliable. For example, one commonly taught generalization is the "silent e" generalization. According to this generalization, "When there are two vowels (i.e., the letters *a, e, i, o, u*), one of which is a final *e*, the first vowel is long and the *e* is silent, or, as some versions put it, "the first one says its name" (as in the print words *smile, bone,* and *cake*). However, in approximately one third of the print words with a silent *e* at the end of the word, the letter representing the vowel is not "long" (as the *o* in *done* and the *a* in *have*). Another commonly taught

generalization is the generalization that "When two vowels occur together, the first vowel is pronounced as a long vowel and the second vowel is silent." This generalization works in approximately half the words where "two vowels occur together" (as in the print words *bead* and *seat*) but doesn't work in the other half (as in the print words *chief* and *been*).

In their search to develop reliable letter–phoneme generalizations, Betty Berdiansky and her colleagues found, among one- and two-syllable words within children's listening vocabularies, more than 211 letter–phoneme correspondences that apply to five or more words.

Letter–Onset and Letter–Rime Correspondences

The relationship between letter–onset and letter–rime correspondences is also complex. There are as many—if not more—letter–onset/letter–rime correspondences as there are letter–phoneme correspondences. As in the case of letter–phoneme correspondences, letter–onset/letter–rime correspondences can represent different sounds (as the letter *g* in the print words *go* and *giant* and the letters *eak* in the print words *beak* and *break*). Regardless of whether one is analyzing sounds into onsets and rimes or into phonemes, the pronunciation of homographs still varies by their grammatical function, words that are spelled morphologically are still spelled morphologically, and the pronunciation of words still varies by dialect.

Nevertheless, letter–onset/letter–rime correspondences are less complex than letter–phoneme correspondence: when there is more than one phoneme in an onset or a rime, there are fewer onsets and rimes per spoken word than phonemes. Most words have more than one phoneme in the onset or the rime in at least one syllable.

Children who have begun to read make analogies between familiar and unfamiliar print words to pronounce unfamiliar print words, and they make these analogies at the onset–rime level. For example, a child who has learned to read the print words *smile* and *small* and has also learned to read the print words *cart* and *part* can figure out that the letters *sm* are pronounced /sm/ and the letters *art* are pronounced /art/. Then, when she encounters the print word *smart*, she is able to pronounce it using her knowledge of letter–onset and letter–rime correspondences in print words she already recognizes.

Phonics Instruction
Out of Context Versus In Context

Since the first century A.D., educators have first taught children letters and/or letter–sound correspondences out of context and then asked children to read words and/or text. However, with the discovery in the 1960s that early readers read words better in context, as in stories, than in isolation, as in lists or on flash cards, some educators have shifted to teaching phonics in the context of text that children have already learned to read.

In this approach, teachers first teach children to read text through instructional techniques such as the language experience approach or shared reading. In these techniques, the teacher points to the print in full view of the children as he or she reads to them. Once the children have learned to read the text, the teacher then teaches them letter–sound correspondences in words in the text.

Children taught phonics in the context of learning to read meaningful stories are more successful in using phonics to figure out new print words than children taught phonics out of the context of learning to read meaningful stories. This can be explained by the fact that when readers read in context they use multiple founts of knowledge to read, not just one fount. They simultaneously use (a) their knowledge of the language represented in the text, (b) their background knowledge on the topic of the text, and (c) their knowledge of letter–sound correspondences to read. Children learning to read bring a competency in spoken language to learning to read that they don't yet have in letter–sound correspondences, and they use this language competency, along with their developing phonic awareness, to read.

Implications

English letter–sound correspondences are complex. However, instruction in letter–onset/letter–rime correspondences has two advantages over instruction in letter–phoneme correspondences: (1) Children who have not yet learned to read can analyze spoken words into onsets and rimes when they cannot analyze them into phonemes, and (2) there are fewer onsets and rimes than phonemes in most words.

Children who have begun to read use their knowledge of letter–onset and letter–rimes correspondences in familiar print words to figure out unfamiliar print words. Children who are taught phonics in the context of learning to read meaningful stories are more successful in figuring out new print words than children who are taught phonics out of context.

Margaret Moustafa

See also Cognitive Development and School Readiness; Literacy; Reading Comprehension Strategies; School Readiness; Spelling

Further Readings

Berdiansky, B., Cronnell, B., & Koehler, J. (1969). *Spelling-sound relations and primary form-class descriptions for speech comprehension vocabularies of 6–9 year olds* (Tech. Rep. No. 15). Los Alamitos, CA: Southwest Regional Laboratory for Educational Research and Development.

Cantrell, S. C. (1999). Effective teaching and literacy learning: A look inside primary classrooms. *The Reading Teacher*, *52*, 370–378.

Clymer, T. (1963). The utility of phonic generalizations in the primary grades. *The Reading Teacher*, *16*, 252–258.

Freppon, P. (1991). Children's concepts of the nature and purpose of reading in different instructional settings. *Journal of Reading Behavior*, *23*, 139–163.

Goswami, U., & Mead, F. (1991). Onset and rime awareness and analogies in reading. *Reading Research Quarterly*, *27*, 150–162.

Moustafa, M. (1995). Children's productive phonological recoding. *Reading Research Quarterly*, *30*, 464–476.

Scholes, R. J. (1998). The case against phonemic awareness. *Journal of Research in Reading*, *21*, 177–189.

Treiman, R. (1983). Onsets and rimes as units of spoken syllables: Evidence from children. *Journal of Experimental Psychology 39*, 161–181.

PHYSICAL DEVELOPMENT

Broadly stated, *physical development* refers to the elements and processes involved in the expression of the biological potential of humans at multiple levels of organization: microbiological, cellular, anatomical, behavioral, and functional. And, inasmuch as humans are material, physical development is at the core of human conative (directed thought and behavior), affective (emotion), and cognitive (mental processing) functioning and development. For these reasons, appreciation and study of physical development is necessary for the advancement of any discipline

concerned with human behavior, such as educational psychology.

Identifying work from a variety of disciplines—anatomy, physiology, evolutionary biology, physics, genetics, chemistry, the cognitive sciences—what follows in the next sections are a brief history of physical development as an element in the study of human development and a review of the systems, including discussion of issues amenable by students and practitioners of educational psychology.

Conceptual Foundations

For most of the 20th century, views of human development were conceptually driven by Cartesian splits, wherein proponents of one position or another jostle for supremacy. These either/or options of what drives human development include continuity/discontinuity, constancy/change, nature/nurture, and stability/instability. Of these, physical systems are most often discussed in terms of nature/nurture. In debating the primacy of nature (biology) versus nurture (environment), physical systems of development are ascribed to the nature end of the argument. Theoretical positions favoring a passive view of physical systems generally reflect a mechanistic viewpoint. Here, environment was viewed as the driving force behind human development, and the organism was seen as reactive, as evidenced in the work of Watson and Skinner. Organismic theorists—early Hall, Gessell, Piaget—suggested that physical systems played an active role in human development. Development was seen as epigenetic—an unfolding of human nature via maturation. At the extreme end of this argument, biological processes are primarily involved in activating the cascade of activities whose teleological end is "humanity," an idea best expressed by the recapitulation that phylogeny (history of the species *Homo sapiens sapiens*) is seen in ontogeny (embryonic history of an individual).

Modern conceptions of human development assert a more contextual orientation, said to result from interactions among physical, cognitive, affective, and conative systems and the environment. Contrary to earlier perspectives, systems are seen to depend on other systems for expression. Such interdependence makes the isolated discussion of any element of human experience problematic. Equally important is recognition that systems exist temporally; development is a fluid process and "snapshot" observations

are often misleading. Among research communities investigating physical development, two assumptions frame data interpretation: the laws of thermodynamics and theories of human evolution. Physicists remind us that the organization of all matter—including humans—follows the laws of thermodynamics. First, energy—such as in cells—cannot be destroyed, only transferred to another source. Next, all physical systems move toward states of entropy or maximum diffusion. Evolutionary theories suggest distal mechanisms for how *Homo sapiens* are structured and how these structures are passed from generation to generation among populations.

Though Darwin is universally linked with evolution, he was certainly not the first to consider it (i.e., Lamarck, Saint-Hilaire). He was, however, the first to express evolution as operating through the generation of population variance among traits and natural selection of those traits in a survival of the fittest. After 150 years of often visceral debate, the central tenants of evolution by natural selection have remained unchanged: Species are capable of overproducing offspring, and resources for support of offspring are limited. Therefore, a struggle for existence among individuals ensues. And, if individuals differ in traits (via adaptations, co-opted outcomes, or evolutionary noise) that enable them to survive and reproduce, and at least some variation in these traits is heritable, different populations will produce, to some degree, different offspring. The how of heritability took longer to fathom. Not until Hamilton's treatise on the evolution of social behavior did the explanatory power of genes enter the discussion of how traits are passed from one generation to the next. Prior to this, heredity was believed to operate at an individual or group level; species evolve because the fittest individuals or groups have a greater potential to reproduce. Instead, Hamilton's work demonstrated that, ultimately, evolution does not occur through individual or group survival. It occurs through survival of the fittest genes.

Cellular and Anatomic Systems Basics

How genes are evidenced brings this discussion back to physics. Like other forms of matter, the basic chemical building block of humans is the atom. Complex groups of atoms, or molecules, conjoin to form the simplest of substances, many of which are nucleic acids. Nucleic acids are organic molecules involved in the genetic transmission of traits. Two types are

identified: deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). Genetic material in each cell is made of segments of DNA called genes. Genes regulate protein synthesis in cells and determine which traits are inherited. RNA carries information from the genes that direct the construction of amino acids into proteins, determinants of the physical and chemical properties of cells. In all, researchers have identified approximately 200 different cell types (e.g., nerve, red blood, epithelial, sperm) among the 100 trillion cells in the average adult human.

Groups of cells and related materials form tissues of which there are four basic types—connective, epithelial, muscle, and nervous. In turn, the joining of two or more tissue types forms organs. Related organs with shared function work in tandem to form the eleven anatomic systems of the human body. Systems include the integumentary (skin, oil glands); skeletal (bones, joints, cartilages); muscular; nervous (brain, spinal cord, nerves); endocrine (hormonal glands); cardiovascular (heart, blood, blood vessels); lymphatic and immune (tonsils, spleen, lymph nodes); respiratory (lungs); digestive (esophagus, stomach, intestines); urinary (kidneys, bladder, urethra); and reproductive (gonads, vagina, penis) systems. In terms of cognitive, affective, and conative functioning, the nervous system is primarily important. As the seat of sensory processes and functions, the brain and related tissues form the ontological center for the human experience.

Developmental Benchmarks

For organization's sake, the listed developmental benchmarks reflect reflex, perceptual, and gross/fine motor skill development and are identified in the age periods in which they occur, on average. Use of age periods is not to be construed as an endorsement of a discontinuity perspective of physical development over a continuity perspective nor should discussion of the "average individual" conceal variation among individuals. Periods and the ages covered include: prenatal (fertilization–birth), infancy (0–2 years old), early to late childhood (2–11), adolescence (11–22 years old), early to late adulthood (22–65+ years). As it is both foundational and of interest for the study of K–12 settings, emphasis is on earlier periods of development.

Recall the assumption that the form and function of the human body has been shaped for millennia via natural selection. Complicating this idea is that every fertile human that has ever lived has either been biologically male or female. Throughout evolutionary history, selection pressures have differed to some degree for males and females; greatest among these are the conditions that gave rise to the traits that ensure conception, birth, and survival of offspring. Downstream from these traits are secondary sexual characteristics, many of which are evidenced during puberty for males (i.e., spermarche, facial hair) and females (menarche, widening hips); other characteristics are evident throughout the school years and include physical performance abilities and risk-taking behavior differences between the sexes. All of this serves as a reminder of the fluidity among the physical systems of development (microbiological, cellular, anatomical, behavioral, functional) when observing or estimating developmental benchmarks.

Prenatal (Fertilization–Birth)

The prenatal period spans from fertilization to birth and is divided into the embryonic and fetal periods. The embryonic period covers the first two months of gestation. After fertilization, the zygote quickly divides into a collection of cells called a blastocyst and covered by another layer of largely nutritive cells, the trophoblast. Approximately a week after fertilization, the blastocyst is implanted into the uterine wall. Within 2 weeks, the blastocyst is differentiated into the ectoderm, mesoderm, and endoderm. Eventually, the ectoderm becomes the lining of the intestinal and respiratory tracts, whereas the mesoderm and ectoderm become bone, muscle, or connective tissues, and the skin or nervous systems, in a process termed *organogenesis*. The beginning of the third month of prenatal development marks the start of the fetal period. Lasting until birth at around the 38th week of gestation, the fetal period is primarily characterized by continued development of the organs.

Infancy (0–2 Years of Age)

At birth, physical development is the measure of the individual, literally. Health care providers throughout the world routinely use two measures of physical performance to determine the fitness of the child at birth: the Apgar index and the Brazelton Neonatal Behavioral Assessment Scale. Whereas the Apgar is used to measure body color, heart rate,

muscle tone, reflex irritability, and respiratory effort, the Brazelton is used to measure neurological development, including response to people and reflex stimuli. Reliance on such tests as the Apgar and Brazelton are reminders of the relevance of physical development for development in general. That the attainment of physical development benchmarks is the best immediate predictor of future cognitive and affective potential highlights the complex relationship among the different systems of human development.

In normally functioning infants, several reflexes are evidenced, each with a specific survival purpose. The sucking and rooting reflexes, for example, aid infants in securing nourishment, whereas the Moro, grasping, and Babinski reflexes are implicated in bringing infants back to equilibrium after averse stimuli (e.g., falling, loud noise). Though all of these reflexes disappear by the end of the first year—some just months after birth—other reflexes, such as breathing or blinking, remain for life. In all instances, the newborn's reflexes provide a scaffold for the development of voluntary behavioral responses. This is not to say that a newborn passively responds to his or her surroundings. Even at birth, sensory perception and largely undetectable motor skills allow individuals to selectively attend to environments.

Of the sensory capacities of newborns, only visual acuity and, to a lesser degree, hearing are not equal to that of adults. Taste, smell, and touch are fully functioning at birth. And, though newborns' eyesight has been estimated to be between 20/400 and 20/800 and they cannot hear many of the softer sounds adults can, by age 2, infants' sensory abilities equal those of adults, including the ability to synthesize information from multiple modalities. Sensory perception is complemented by emerging motor skills. During the first year, infants rely on gross motor skills, with fine motor skills improving by age two. Throughout infancy—and beyond—motor skill development reflects two trends: cephalocaudal (from the head down) and proximodistal (from torso to limb extremities). The cephalocaudal trend is evidenced, for example, in the transition from ability to lie in a prone position with head up (1 month of age) to using arms for support (3 months of age) to sitting with support (6 months of age) to standing with support (8 months) to, eventually, standing alone (11 months). Optimal conditions for physical development of infants include extended amounts of sleep (>15 hours/day) and a calorie-rich diet (twice the calories per pound compared with adult intake). The

results are marked. By the end of the first year, the average infant's weight triples and length increases by 50%.

Early to Late Childhood (2–11 Years of Age)

Following infancy, the pace at which physical development occurs during childhood seems quite a bit slower. In many ways, this is true. The brain never grows faster than it does during infancy. Intermodal sensory perception is complete in the first year. And, regarding motor skills, for example, the rapid progression from being unable to control one's movements to the ability to walk independently is followed by the relatively long amount of time needed to progress from initial use of a pincer grasp (sometime during the first 12 months) to adult-level hand and finger control (10 years of age). In other ways, physical development continues apace; we are never more active than we are as children, and yearly height and weight gains tend to be consistent from infancy throughout childhood.

Adolescence (11–22 Years of Age)

Reproductive maturity is a hallmark of adolescence. By the end of puberty, males produce sperm; females produce ova. Internal processes of gamete production are announced to the world via secondary sexual characteristics. Male characteristics include an increase in height; pubic, facial, and body hair growth; deepening of voice; increased muscle mass; broadening of the shoulders; penile growth; and spontaneous erections. Female characteristics are height increases, breast enlargement, widening of hips, pubic and body hair growth, and deepening of voice. Though the ages in which males and females enter puberty tend to differ (males = 12½ years of age; females = 10½ years of age), changes in hormonal levels are implicated in both. All humans produce estrogen and testosterone. During puberty, males demonstrate a significantly greater increase in testosterone production compared to females, and females produce significantly more estrogen than males.

Early to Late Adulthood (22–65+ Years of Age)

By their early 20s, most people have reached a maximum height. For many of us, weight continues to increase, but this is a function of a sedentary lifestyle

and poor nutrition rather than maturation. That said, signs of aging reflect the second law of thermodynamics and are demonstrated in every physical system. The heart muscle, for example, becomes less efficient over time. Elasticity lessens, and fatty deposits tend to form on the walls of arteries, which makes the heart work harder than when younger, often leading to hypertension. More apparent examples include side effects of menopause, osteoporosis, arthritis, and a general weakening of sensory perception potential.

Physical Development Crises in U.S. Public Education

Since the inception of the normal school, the focus of teacher education programs has been on the training of candidates in specific curricular theories and pedagogy. Modern accountability movements have strengthened the tendency. Most early childhood and physical education teacher training excluded, teacher education often omits extended consideration of how optimal physical development is related to learning and motivation. Often, an educator's first exposure to the deleterious effects of a crisis of a physical system— that is, poor nutrition, eating disorders, substance abuse, sexually transmitted diseases—is after training has been completed. Traditionally, students with these problems are referred to support staff, such as school nurses, guidance counselors, and coaches. But K–12 teachers revisit the effects every day in their classrooms. Here, the work of educational psychologists, among others, can be instructive.

Nutrition, Exercise, and Body Image

The World Health Organization (WHO) identifies a body mass index (BMI) greater than 25 kg/m^2 as overweight. The United States leads the world in the number of children who are overweight or at risk for being overweight ($> 30\%$), yet rates are high around the world, including Australia, China, Eastern Europe, the Middle East, Pacific Islands, and the United Kingdom. Recent estimates suggest levels in the identified regions have nearly tripled compared to those of 25 years earlier. These children face increased risk for bone and joint problems, cardiovascular disease, type 2 diabetes, dyslipidemia, and asthma, in addition to academic and social discrimination. And, as public schools are among the larger food service providers to children, K–12 settings have a significant role in

this problem. A long-standing response has been to provide adequate nutrition education for students. To date, all public schools offer some form of nutrition curriculum, but elementary-school students receive the most attention in this area.

Though childhood obesity is a serious medical problem, most children with physically active lifestyles can consume copious amounts of sugars, unsaturated fats, and so forth, with few short-term effects. Child and adolescent basal metabolism rates (BMR) are higher than adult BMRs and, when coupled with an active lifestyle, children can usually maintain a healthy weight. Historically, public education has included physical education as part of the K–12 curricula, again with a focus on the elementary-school years. Now only about half of students attend physical education classes at least once a week and only a third attend daily. Recent trends show a decrease in the amount of time students are engaged in physical activity during school hours. This pattern means that less than a third of high-school students get recommended levels of physical activity (i.e., exertion resulting in increased heart rate and respiration for at least 60 minutes a day for 5 of 7 days of the week). Add this to researchers' findings that school-age children spend approximately three hours a day using multimedia outside of school, and it is not surprising that most children develop inactive lifestyles as adults. Willingly or not, K–12 settings play a significant role in a modern health crisis.

The other end of the continuum is seen in the growing number of preteens and adolescents who suffer from body image disorders. As much as 10% of U.S. females suffer from such disorders some time during their lives. Males suffer at a far lower rate. Whereas most K–12 students have a healthy concern for body shape and size, for others, attention to body image is a life-threatening addiction. Disorders include anorexia (starvation), bulimia (binge and purge), or bigorexia (steroid use). In all instances, the health of students is in imminent danger if disorders go unchecked. Educating teachers about the warning signs is imperative, because, more often than not, these students are socially competent high achievers and, thus, are less likely to draw concern than other students.

Substance Use

Worldwide substance use varies, but the health risks for substance users are consistently documented. Because of their underdeveloped bodies, substance

use among school-age populations causes an even greater risk than for older populations. Among those students who use illegal substances, the overwhelming majority are adolescents, who are more likely to use illegal substances as adults than are their drug-abstinent peers. To this end, drug education curricula are in place in K–12 settings and supported by both government and nongovernment agencies. In the United States, for example, drug use prevention programs are aided through the efforts of Drug Abuse Resistance Education (D.A.R.E.), which is offered by local law enforcement, and programs offered by the Office of National Drug Control Policy (ONDCP). Evaluation of the effectiveness of drug prevention programs is mixed, but it is evident that a large portion of teenagers worldwide have experience with substance use. And, though much of the drug use is considered experimental, the potential for long-term health risks causes alarm among educators.

Drug use among students usually refers to alcohol, tobacco, or marijuana. Among these, alcohol is, far and away, the most used drug by school-age students in the United States and Europe. In a recent study of U.S. students, researchers suggested that almost 3 in 4 high-school students have tried alcohol and close to 2 in 4 have tried it in a 30-day period. One in four students has engaged in binge drinking (five or more drinks) in a given month. Tobacco is another commonly used drug among students, especially high-school students. More than half admitted to having tried tobacco (cigarettes, cigar, and smokeless), and more than a quarter report being current users. Alcohol and tobacco use among European teenagers is even higher, yet American teens seem to use drugs such as marijuana, cocaine, acid, and Ecstasy at higher rates than their European peers. In the United States, research suggests approximately 40% of high-school students reported smoking marijuana, and 1 in 5 had smoked at least once in a 30-day period. Other drugs are consumed in significantly lower quantities. Drugs and use among students at least once in their lives include inhalants (12%), cocaine (8%), hallucinogens (9%), methamphetamines (6%), MDMA (6%), and heroin (2%).

Sexual Activity, Pregnancy, and Sexually Transmitted Diseases

Researchers have consistently documented that worldwide at least 50% of school-age children have had sexual intercourse by the end of their last year of compulsory education. Some estimates suggest the rate is as high as 70%. Yet, despite the relative parity in activity level across the globe, the United States has double the teenage pregnancy rate of other industrialized countries; the United Kingdom is second on the same list. Socially and academically the consequences of teenage pregnancy are significant. More than half of teenage mothers end up on government assistance, and less than a third of teenage mothers complete their compulsory education. Condom usage is lower among teens in the United States and United Kingdom than it is in comparable countries. Both facts, researchers suggest, are related to greater reliance on abstinence-only education in the United States and the United Kingdom than in countries that provide sex education. More comprehensive sex education (curricula including topics of birth control, masturbation, or parenting) has been the norm in many Asian (e.g., India, Mongolia, South Korea, Sri Lanka) and European (e.g., France, Germany, Sweden) countries since the 1970s. To a lesser degree, more comprehensive curricula include information on lesbian, gay, bisexual, and transgender orientations and types of intercourse.

With a higher rate than pregnancy, sexually transmitted diseases (STDs) wreak havoc on the physical—not to mention affective and cognitive—development of school-age children. Educators are in a double bind when it comes to STD prevention. STD symptoms are often mild or absent for years after an infection is contracted so can be misinterpreted by adolescents and adults as no cause for concern. But, if left untreated, STDs can be life-threatening; the most serious STD is HIV/AIDS. Opportunistic infections and myriad cancers that affect nearly every physical system of the human body are symptomatic of advanced HIV infection. Use of antiretroviral therapy has been shown to extend life postinfection by more than five years. Of course, as the viruses and treatment options evolve, survival estimates will change. Bacterial infections such as chlamydia, gonorrhea, and syphilis also have health consequences if left untreated. Unlike HIV infection, some bacterial infections are evidenced differently for males and females. In males, for example, chlamydia shows little effect, but in women it is associated with pelvic inflammatory disease and, in some cases, infertility. Gonorrhea affects females in ways similar to chlamydia; in males, it can cause epididymitis, which can also lead to infertility. More commonly, school-age children

contract human papilloma viruses and genital herpes. Viral infections result in less debilitating conditions than other STDs; symptoms include genital sores, flu-like symptoms (herpes), and warts (HPV). However, presence of HPV or herpes increases an individual's susceptibility and transmission of HIV. Like other disturbances of physical systems in school-age children, STDs disproportionately affect specific cultural groups. In the United States, for example, African American and Native American females have significantly higher STD rates than any other ethnic/gender grouping. In all, sex education—including abstinence only—has been shown to work. Exposure to sex education curricula is associated with decreased STD rates and lower teenage pregnancy rates. Recent success, for example, in slowing the AIDS epidemic in Africa is attributed to the educational efforts of the United Nations and nongovernment organizations such as the WHO.

Future Directions

In the past century, educational psychology has emerged as a field that contributes greatly to a collective understanding of the behavior of school-age children. Though the roots of educational psychology as a field of study emerged from basic principles of physical development, interest in the topic waxes and wanes. Toward the middle and end of the 20th century, research in educational psychology reflected the assumption that human experience is largely a recapitulation of an infinitely plastic culture. Unfortunately, a concern over essentialism or nativism turned into a disdain for research suggestive of a role for the physical in cognitive, affective, and conative development. More recently, there has been a gradual reawakening of the role of physical development in the human experience via fields such as the cognitive sciences and evolutionary biology. In all things, educational psychology is best informed when all systems of development are considered in tandem.

Sean Alan Forbes

See also Abstinence Education; Drug Abuse; Eating Disorders; HIV/AIDS; Malnutrition and Development

Further Readings

Alexander, P. A., & Winne, P. H. (Eds.). (2006). *Handbook of educational psychology* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.

Bornstein, M. H., & Lamb, M. E. (Eds.). (2005). *Developmental science: An advanced textbook* (5th ed.). Mahwah, NJ: Lawrence Erlbaum.

Pinker, S. (2002). *The blank slate: The modern denial of human nature.* New York: Viking Penguin.

Tortora, G. J., & Derrickson, B. H. (2006). *Introduction to the human body: The essentials of anatomy and physiology.* Hoboken, NJ: Wiley.

PIAGET'S THEORY OF COGNITIVE DEVELOPMENT

To properly understand Jean Piaget's theory of cognitive development, it is important to consider it within the larger context of his work. Although Piaget is recognized as one of the greatest developmental psychologists, he described his own work as "genetic epistemology." Genetic (Greek *genno* = give birth) here refers to the origin and development of knowledge, rather than to genes, as the word is used today. The main goal of Piaget's epistemology was to explain the generativity and rigor of human knowledge. *Generativity* refers to novelty and invention, whereas *rigor* refers to logical necessity, that is, that an answer must necessarily and logically be true, could not be otherwise, and must universally hold to be true for all rational persons.

Even though Piaget also approached the genesis of knowledge from the perspectives of phylogeny and the history of science, the major portion of his work addressed this issue by studying the development of knowledge in children. In this way, Piaget addressed fundamental epistemological questions about the origin, development, and validity of knowledge in general. He concluded that the development of knowledge is a constructive process, and he emphasized the child's active role in the construction of knowledge: Knowledge is constructed through a process of active exchange between the individual and his or her environment. Piaget's constructivist theory is essentially a theory of dynamic self-organization, which is rooted in biological functioning, with cognitive development representing the extension and continuation of this process of biological self-organization to a new level of functioning. This process of cognitive development results in the construction of increasingly advanced forms of thinking that Piaget described as progressing through a series of stages.

Piaget's Constructivist View of Knowledge and Development

Theories of development are based on views of the nature of knowledge, and therefore, Piaget argued that it is essential to examine foundational assumptions about the nature of knowledge. He argued against "copy theories" of knowledge, according to which knowledge consists of acquiring images, pictures, or representations that match reality. A flaw in these theories is that it is not possible to check the accuracy of such copies except by comparing them to reality itself. But such comparisons are not possible according to copy theories, because the point of the copy was to provide knowledge of reality; if we could directly access reality in order to compare our representations against it, we would not need such representations in the first place. Therefore, this view does not explain the development of knowledge about the world; instead, it already presupposes its existence.

For Piaget, knowledge, rather than consisting of images or representations, is built up through action on the world and through coming to know what can be done to aspects of the world. Acquiring knowledge through action begins in infancy with simple acts such as pushing and pulling, and continues throughout development, because, according to Piaget, even the most sophisticated forms of thought are interiorized actions, now carried out mentally. Knowledge is not innately preexisting within the child, nor does it arise solely from empirical experience with objects, such that this experience produces a simple copy of the object. Rather, the essential characteristic of Piaget's constructivism is that intelligence is constructed through the child's continuous interaction with the world. In this sense, Piaget considered his constructivism a third way that avoids the problems with both nativism and empiricism.

Central in describing the process of development are Piaget's concepts of *scheme, assimilation*, and *accommodation.* These concepts describe the functional relation between the individual and the world at any point in development. A scheme is a general structure that is applied in a particular situation, and it is that aspect of any activity that can be generalized. For example, at the sensorimotor level, schemes are general patterns of activity that can be repeated, such as the sucking scheme, which can be applied to different objects. Assimilation is the incorporation of objects or events into an already acquired pattern of activity, or

scheme. By integrating objects and events into preexisting knowledge structures or schemes, assimilation gives them meaning (e.g., "suckable"). Infants not only suck on nipples, but also on fingers; that is, a finger may be assimilated to the sucking scheme. But because of differences in the experience of sucking a finger (e.g., it provides no nourishment), the infant differentiates this experience, and *accommodation* is said to have occurred: Patterns of activity differentiate to allow for the assimilation of novelty. Assimilation and accommodation are inseparable and describe two fundamental aspects of any activity in the process of adapting to the world, that is, acquiring knowledge.

The concepts of assimilation and accommodation have several implications. First, they express the idea that development is a continuous process that, at the same time, leads to structural change (differentiation and integration of knowledge structures). Second, activity is always organized in the sense that it is based on a structure (otherwise objects interacted with would be devoid of meaning). Structures, however, do not exist as an entity in the mind that results in reasoning; rather, they exist as potential coordinations of operations. Third, assimilation and accommodation continue, on a functional plane, the material process of self-organization (metabolism), thereby securing the continuity between biological and psychological functioning.

Based on this constructivist view of knowledge, Piaget described a series of stages, or forms of thought, in the development of intelligence. These stages build on each other and, therefore, necessarily develop in the following sequence.

Sensorimotor Intelligence

During the first stage of cognitive development, infants interact with the world through sensorimotor patterns of activity that gradually come to be differentiated and coordinated, as a result of interaction with the world. Because of the relative lack of differentiation and integration of action patterns, infants' experience of the world is undifferentiated from and fused with their own activity on the world. Piaget argued that infants' initial experience of the world is centered on their own bodies, which he referred to as *egocentrism.* This does not mean that infants are focused on themselves (self-centered), but rather that they have not yet constructed an understanding of themselves as objects existing among other independent objects.

Over the first one and half to two years of life, this initial egocentrism, or centration on the self as the reference point of epistemic experience, is gradually overcome, thanks to the functional interplay of assimilation and accommodation. Piaget described this process as occurring over a series of six substages. To get a sense of how radical Piaget's theory is, consider that it is during this period that, according to Piaget, the infant gradually constructs a sensorimotor, practical understanding of what will only later be reflectively understood as space, time, causality, and objects. For adults, such a conception of the external world is simply taken for granted, and assumed to be given by perception. Yet, according to Piaget, infants must gradually construct such an understanding of the external world.

Particular interest was generated by Piaget's description of infants' development of object permanence; that is, infants' growing understanding that objects exist as "things out there," independent of their own activity. During the sensorimotor period, object permanence undergoes a systematic, stagewise development. For example, at substage 3 in the development of object permanence, infants will not search for an object if it is completely covered, but they can retrieve an object if part of it is still visible. At substage 4, infants can successfully search for an object even it is completely covered. However, if they have found it under cover A, and then they see it placed under cover B, they will still continue to search for it under A. This curious phenomenon, referred to as the *A not B error*, has generated a great deal of research. Piaget's explanation for this characteristic error is that the object is not yet sufficiently separated from the infant's own action of finding it in the first location. With increasing integration and combinations of schemes, the object will eventually be conceived of as external to and separate from the infant's own activity: The more an infant can do with an object (e.g., grasp, suck, look, drop), and the better they can coordinate these action schemes, the more the object takes on an existence independent of the infant's activity.

Sensorimotor intelligence is practical or lived knowledge, which means that such knowledge is dependent on interaction with objects and is not yet reflective in nature. Further development requires a gradual process in which this knowledge is conceptualized and reconstructed at a higher level within the organization of the succeeding stage. The sensorimotor stage ends with the emergence of the symbolic or semiotic function, which is the ability to use symbols or signs to represent objects or events that are not present. The semiotic function is made possible by the interiorization of imitative actions such that these actions are performed internally and serve as images for the signification of schemes.

Preoperational Intelligence

Thanks to the emergence of semiotic functioning, cognition at the preoperational stage is no longer limited to the immediate here and now, but can re-present objects that are not in the immediate spatio-temporal field. As a result, the interplay between assimilation and accommodation becomes more complex because it involves both perceptual and representational levels of functioning, and thus absent objects as well as present objects. The development of the semiotic function requires that the child laboriously reconstruct, on the new representational plane, the practical concepts of object, space, causality, and time that had been constructed and only practically understood at the sensorimotor stage; Piaget termed this process of reconstructing concepts at a qualitatively different plane *vertical decalage*.

The semiotic function manifests itself in a number of different activities such as deferred imitation (i.e., imitation in the absence of the model), pretend play, drawing, psychological functions based on mental images (e.g., recall memory), and language. These activities are practiced and refined during the first substage of the preoperational stage, which Piaget termed *preconceptual thought*. Preconceptual thought is no longer tied to particular objects or events, but it fails to distinguish between individual members of a concept and the generality of concepts. In the second substage of preoperational thought, lasting from about 4 to 6 or 7 years and termed *intuitive thought* by Piaget, representational schemes become increasingly coordinated, and children become capable of relating two representational schemes to each other by means of a unidirectional logical relation. For example, thought and attention may be centered on one dimension, such as using height in order to infer amount of liquid, even though neglecting the width of the container. However, intuitive thought is incapable of understanding the simultaneous reversible and bidirectional nature inherent to logical operations.

Concrete Operational Intelligence

During the next stage of development, concrete operational intelligence, the semiotic functions of the

preoperational level come to be coordinated into what Piaget called operational systems. Operations are actions (e.g., putting like objects together, putting objects into one-to-one correspondence) that are interiorized and reversible. Reversibility means that transformations that have occurred in reality can be compensated for on the representational plane by incorporating these transformations into a system of logical relations. As a result, the child at the level of concrete-operations is liberated from centering on only one aspect of a situation (e.g., a perceptually salient aspect of situation).

Piaget constructed a number of tasks to assess concrete operational thought. Many of these entail forms of conservation. Conservation refers to the understanding that a whole remains intact despite undergoing transformations. For example, the number of objects in a collection, as a whole, does not change if the objects are rearranged. An operative understanding of conservation, therefore, is logical in nature—it is a logical truth that is not given by empirical observation of transformations. Piaget's conservation tasks were designed such that children who lacked an operative understanding of conservation would be misled by the appearance of transformation in the tasks. For example, conversion of liquid requires understanding that the amount of liquid does not change even if its shape is transformed by being poured into a tall thin glass.

A limitation, however, of concrete operational thinking is that although it is logical, it is still restricted to reasoning about actual objects. For example, consider an experiment in which a child is faced with an experimenter holding a poker chip in his hand saying, "Either the chip in my hand is green or it is not green." The child is then asked if the statement is true or false, or if he or she cannot tell. The child using concrete operational reasoning would be uncertain and would have to ask to see the chip. However, the statement is a tautology; it has to be true regardless of the chip in the experimenter's hand. That is, the statement is necessarily true based on its *form*, and therefore, looking at the actual chip is not required.

Formal Operational Intelligence

The next stage of development, formal operational intelligence, involves operating on logical classes or forms rather than on concrete objects, which are specific instantiations of logical classes. Piaget considered hypothetico-deductive reasoning to be the

hallmark of formal operational thinking. This form of thinking involves the reversal of the direction between reality and possibility: Whereas on the level of concrete operations, possibility remains an extension of reality, on the level of formal operations, reality is subordinated to possibility. As a consequence, the adolescent can now reason about possibilities. Bärbel Inhelder and Piaget studied the emergence of formal operations by presenting children and adolescents with problems involving concrete material to be manipulated in order to discover scientific laws. For example, the pendulum task involves discovering which of several factors (length of string, weight of object, height of dropping point, or force of push) determines the frequency of the pendulum's oscillations. These experiments revealed that children differed qualitatively in their approach to scientific problems compared to adolescents. Although children were capable of classifying and cross-classifying an independent variable along one dimension, and of putting these seriations into correspondence with their effects on the dependent variable, they still failed to design systematic experiments, and, as a result, did not supply adequate proof for their statements. By contrast, adolescents formulated hypotheses and tested them systematically by controlling all variables except the one under investigation (isolation of variables) in order to gradually converge on the correct hypothesis.

Piaget recognized that there is more to thinking than logic, but logic and formal operational thinking about possibilities are not separate from social life and are intertwined with adolescents' construction of a scale of values that underlies their plans as they enter adult society. Thus, affective life, for Piaget, is not separate from cognition.

Structure, Equilibrium, and Equilibration

The coordination of operations (interiorized actions) into structures leads to Piaget's solution to the problems of generativity and rigor of human thought. Generativity of thought is due to the coordinations of actions, resulting in a range of new possibilities. Rigor, or necessity—understanding that an answer is necessarily correct—follows from the completion of a *structure*, which then entails *logical necessity*. For example, a child who fully understands the concept of numbers knows that 5 plus 7 is necessarily 12. The *equilibrium* and closure of a structure results from the

operations being *reversible;* every operation, such as addition, can be compensated for by another operation, such as subtraction. Higher forms of knowledge involve more adequate forms of equilibrium. Development is a process leading to increasingly more stable (complete and consistent) forms of equilibrium, and, therefore, development is progressive; it is not mere change. Equilibrium can involve a balance among a child's own activities (*organization*) or between the child and the environment (*adaptation*). *Disequilibrium* could be cased by gaps or contradictions in knowledge, and *equilibration* is the process of achieving a new, and often more complete, form of equilibrium following disequilibrium.

Although Piaget's stage theory is the best-known aspect of his work; this has been at the cost of neglecting his theory of equilibration. Perhaps because Piaget was trained as a biologist, it was natural for him to classify children's thinking into different stages or forms of thought. Yet, for Piaget, this was only the first step in understanding the development of knowledge. The second and more important task was to explain development from one form of thinking to another. This was the goal of Piaget's theory of equilibration, which brings out his interactive process account of development. Piaget emphasized equilibration as a process rather than equilibrium as a state. Every point of equilibrium is only partial. Piaget argued that equilibration is an essential factor in development, in addition to maturation and experience with the physical and social environment.

Piaget's later work focused on delineating in more detail the specific processes involved in equilibration. In this work he emphasized the roles of consciousness, affirmation and negation, contradiction, and reflective abstraction. Essentially, Piaget suggested that in the course of their interaction with the environment, and in the context of encountering obstacles to their actions, children become increasingly aware of their knowledge schemes and the coordinations involved in their actions. By reflecting on the coordinations of these actions, children become aware of the coordinatory structure involved in their actions. Reflective abstraction, thus, can be seen as a mechanism that, at each level of knowing, abstracts form (i.e., the coordinatory structure of action) from content and, in turn, projects this form to a higher level. With each new and higher stage, the forms become increasingly abstract. Through the mechanism of reflective abstraction, then, development proceeds by way of successively conceptualizing the forms or structures of knowledge underlying previous knowing levels.

Understanding Piaget's Theory Through Evaluating Common Criticisms

Although Piaget has been extremely influential in developmental psychology, his theory has been misinterpreted in numerous ways, partly because his goals have not always been recognized. There are now two views of Piaget's theory: the familiar "received view" that has become entrenched in textbooks, and a more recent and close reading of Piaget's work advanced by Michael Chapman and others that differs in striking and important ways. From the perspective of the received view, Piaget is acknowledged as a pioneer in many areas but nonetheless is heavily criticized for a number of reasons. The implication is that Piaget has little to offer current research and theory in developmental psychology. Therefore, to give a contemporary summary and assessment of Piaget's theory, it is now necessary to discuss both views of his work.

The primary diversity of interpretation of Piaget's theory revolves around the concept of developmental stages. Piaget's idea of stage is commonly interpreted to imply that once children demonstrate a form of reasoning, such as concrete operational reasoning (e.g., conservation tasks), they are in this stage and therefore should be able to pass all other concrete operational reasoning tasks. That is, the principle of conservation is the same, whether the concept conserved is substance, volume, weight, or liquid. However, there is now overwhelming evidence of inconsistency in reasoning across tasks that have identical logical formal properties, such as conservation. This inconsistency is known as *horizontal decalage.* For example, children's understanding of conservation develops in the following order, with each type being separated by approximately two years: substance (7–8 years), weight (9–10 years), and volume (11–12 years). This evidence is generally seen as a fatal flaw for Piaget's theory. However, Chapman pointed out that Piaget never did make this claim so often attributed to him about consistency in reasoning within stages, and, in fact, Piaget stated the opposite in several of his writings. If we begin from Piaget's basic insight that thought originates in action, then horizontal decalage is not an embarrassing surprise at all; rather, it ought to be expected. That is, as children engage in new forms of actions (e.g., displacing

water, weighing with a scale), new understandings (e.g., types of conservation) should emerge. This discussion also highlights the point that Piaget was classifying forms of reasoning (knowledge) as the object of his study; he was not classifying children as being at one particular stage, as is commonly understood.

It has also been claimed that Piaget underestimated children's abilities because his tasks include extraneous factors (i.e., factors not intrinsically related to the concept being tested). Consequently, researchers have modified Piaget's tasks and removed what they considered to be extraneous factors in order to uncover children's true competence. For example, the conservation of numbers task was administered with two or three objects instead of five objects, and younger children did pass this simplified task. However, this work was later criticized because it became evident that children could pass the simplified tasks with different forms of reasoning. That is, the simplified tasks were no longer assessing the form of reasoning that they were originally designed to assess. Furthermore, this line of research overlooks the crucial point that, within Piaget's theory, age is only an indicator not a criterion for children's competence. The issue that competencies develop in an ordered, sequential manner was more important to Piaget than the question of when these competencies emerge.

Piaget has been criticized for neglecting the importance of social factors and language in development. Moreover, it is generally assumed that he took a strictly individualistic perspective on development. However, in several of his books Piaget emphasized that social interaction is an essential factor in development but that it is necessary to go beyond such obvious statements to clarify how particular forms of social interaction influence development. Early in his career, Piaget argued that reasoning develops from the social process of argumentation. Later, Piaget recognized the roots of logical thought in infants' prelinguistic activity and thus argued that although social factors are necessary, they are not in themselves sufficient as a complete explanation for cognitive development. Social factors are important in knowledge being imparted from one generation to another, but this could not explain how new forms of knowledge emerge, nor how children develop to the point at which they can begin to assimilate such socially available knowledge. Although Piaget did focus on the child's physical action on the world, for a full appreciation of Piaget's thought, his research should be viewed in the context of the larger framework in which he worked.

Implications for Education

Interpretations of Piaget's theory as individualistic might suggest that his theory has little to offer education. However, Piaget himself was explicitly concerned with education, and his theory is a general approach to cognitive development that has implications for social–cognitive development and for education. Piaget's theory directs our attention to the child's level of development because a child can only understand instruction if he or she has developed structures or forms of understanding with which to do so. Furthermore, according to Piaget, knowledge is constructed through interaction with the world, and he emphasized the child's active role in the constructive process. This suggests that rote memorization or passive reception by children is not the best way to learn. However, Piaget's theory does not imply that there is no role for teachers; teachers are essential in creating situations that facilitate children's ability to develop understanding.

Another example of the implications of Piaget's work for education follows from his approach to moral development in which he emphasized the role of two types of relationships: constraint and cooperation. Relationships of constraint involve unilateral respect and the imposition of views from authority. In contrast, relationships of cooperation are best suited for the development of knowledge because they involve mutual respect, and each person is obliged to listen to the other and to fully explain themselves. This situation is most likely to lead to mutual understanding, which is essential in the development of all forms of knowledge.

Jeremy I. M. Carpendale, Ulrich Müller, and Maximilian B. Bibok

See also Conservation; Constructivism; Egocentrism; Equilibration; Object Permanence

Further Readings

Bremner, G. (2001). Cognitive development: Knowledge of the physical world. In G. Bremner & A. Fogel (Eds.), *Blackwell handbook of infant development* (pp. 99–138). Oxford, UK: Blackwell.

Chapman, M. (1988). *Constructive evolution: Origins and development of Piaget's thought.* New York: Cambridge University Press.

Chapman, M. (1992). Equilibration and the dialectics of organization. In H. Beilin & P. B. Pufall (Eds.), *Piaget's theory: Prospects and possibilities* (pp. 39–59). Hillsdale, NJ: Lawrence Erlbaum.

Furth, H. G. (1969). *Piaget and knowledge.* Englewood Cliffs, NJ: Prentice Hall.

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence.* New York: Basic Books. (Original work published 1955)

Lourenço, O., & Machado, A. (1996). In defense of Piaget's theory: A reply to 10 common criticisms. *Psychological Review*, *103*, 143–164.

Piaget, J. (1963). *The origins of intelligence in children.* New York: Norton. (Original work published 1936)

Piaget, J. (1970). Piaget's theory. In P. Mussen (Ed.), *Carmichael's manual of child psychology* (3rd ed., pp. 703–732). New York: Plenum Press.

Piaget, J. (1971). *Biology and knowledge.* University of Chicago Press. (Original work published 1967)

Piaget, J. (1971). *The construction of reality in the child.* New York: Ballantine. (Original work published 1937)

Piaget, J. (1971). *Science of education and the psychology of the child.* New York: Penguin Books. (Original work published 1969)

Piaget, J. (1972). *The principles of genetic epistemology.* London: Routledge & Kegan Paul. (Original work published 1970)

Piaget, J. (1995). *Sociological studies.* London: Routledge. (Original work published 1977)

Piaget, J., & Inhelder, B. (1969). *The psychology of the child.* New York: Basic Books. (Original work published 1966)

Smith, L. (1993). *Necessary knowledge.* Hove, UK: Lawrence Erlbaum.

Smith, L. (2004). Developmental epistemology and education. In J. I. M. Carpendale & U. Müller (Eds.), *Social interaction and the development of knowledge* (pp. 175–194). Mahwah, NJ: Lawrence Erlbaum.

POVERTY

Poverty has historically been defined as some form of expression for material and/or financial hardship that results in the inability to meet basic human needs. The most publicly acceptable versions are focused on income- or resource-based definitions of poverty. These definitions include the following: *absolute poverty* (insufficient income to meet basic human needs) and *relative poverty* (a level of income that does not allow for consumption that meets community standards or goes below a particular percentage threshold of the total population, e.g., 25% or less of median income).

Recent scholars have suggested that *social exclusion* be also included among the definitions of poverty. Social exclusion can be a result of racial/ethnic discrimination, gender discrimination, geographic isolation, educational attainment, or other such sociodemographic characteristics that are associated with social isolation.

Other scholars, such as Robert Chambers, have suggested that there are four clusters of definitions that have defined poverty historically: (1) income-poverty, (2) material lack or want, (3) capability deprivation, and (4) a multidimensional view with material wants as only one of several "mutually reinforcing dimensions" (p. 2).

Income poverty in this categorization relates to the previously mentioned absolute poverty threshold. For material lack or want, the conceptualization relates to not only income poverty but also the ability to secure such needs as clothing, shelter, transportation, and other essential assets. The third cluster of definitions, capability deprivation, relates to the aforementioned social exclusion. The ability of individuals to thrive is compromised as a result of restrictions on what one can and cannot do and what can and cannot be. The fourth cluster takes an ecological, systems-theory approach to poverty that incorporates multiple dimensions and processes.

Theories of Poverty

Ted Bradshaw has outlined what he believes are the five competing theories that drive antipoverty strategies:

1. Individual deficiencies

2. Cultural belief systems that support subcultures of poverty

3. Political–economic distortions

4. Geographical disparities

5. Cumulative and circumstantial origins

In the individual deficiencies framework, adherents suggest that poverty results primarily from a person's personal attributes (e.g., intelligence and diligence) or from poor choices (e.g., dropping out of school). In the individual deficiencies theoretical framework,

poverty is the nexus between individual deficiencies that have been passed down from previous generations and inherent/genetic individual deficits. A suggestion by Richard Herrnstein and Charles Murray that the nexus between race and intelligence are tied to economic well-being has generated significant controversy and has been generally discredited. This set of theories, however, has a long history.

The Virginia Legislature in 1619 ordered any idle able-bodied person be bound over for compulsory labor, and early conceptualizations of poor persons categorized them into two groups: the "deserving" and "undeserving" poor. The undeserving poor were those who were deemed able to work but not currently working. They were believed to be poor as a result of their lack of initiative, intelligence, or some other personal deficit. Deserving poor were those who were poor through circumstances beyond their control; for example, orphans and the elderly.

The second theory, often referred to as the "culture of poverty," suggests that over generations, belief systems, values, and skills have been transmitted to successive generations that perpetuate their poverty. Perhaps the most famous expression of this ideology was given by Senator Daniel Patrick Moynihan in his report from the U.S. Department of Labor titled, "The Negro Family: The Case for National Action," or more popularly known as the "Moynihan Report." The report's main tenet being that the "crumbling" of the "negro" family had substantially contributed to the "cycle of poverty." The report met with scathing criticism from civil rights groups of the time for ignoring the social realities of African Americans at that time.

The third and fourth theories are related in that their adherents frequently reject individual and cultural explanations of poverty in favor of explaining the etiology of poverty through dysfunctional social and economic systems. The third theory, the "political and economic distortions" theory, suggests that the social, political, and economic systems perpetuate poverty. As Bradshaw notes, perhaps the most famous early proponent of this theory was Karl Marx, who believed that capitalism created a permanent unemployed class in an effort to keep wages low. Racial/ethnic, gender, age, and other forms of discrimination that result in low-wage employment or unemployment fall under this theory. The "geographical disparities" theory of poverty suggests that where one lives influences the likelihood of one being in poverty. William Julius Wilson has suggested that for the urban poor,

the flight of well-paying jobs into the suburbs has resulted in a form of economic isolation that creates substantial barriers to gainful employment (e.g., transportation). Those who are able to leave these areas to be closer to jobs, avoid crime, and have their children attend better schools do so, thus leaving behind an increasingly marginalized population. This cycle continues as employers do not place their firms in areas deemed unsafe, where infrastructure in deteriorating, and/or where the availability of well-trained workers is limited.

The final theory in Bradshaw's taxonomy of poverty theories, published in 2006, is the cumulative and circumstantial theory of poverty and is more commonly referred to as the "cycle of poverty." He believes that this is the most intricate of the theories, in that it incorporates the previous four elements in describing a pattern of poverty that is perpetuated by an interlocking set of circumstances and personal characteristics (some driven by circumstances) that results in impoverishment.

Measuring Poverty

The poverty line that has been used in the United States since 1963 was developed by Mollie Orshansky, then working for the U.S. Social Security Administration. Orshansky, previously a home economist, took the amount of income necessary to meet the "economy" food plan as set by the U.S. Agriculture Department for a family of four and multiplied that number by three. The multiplier was derived from a previous survey, the 1955 Agriculture Department's Household Food Consumption Survey, which had determined that the normative expenditures for food for a family of three or more persons equated to a third of all family expenditures. Poverty thresholds based upon differing family sizes were developed by Orshansky in 1965. Currently, the Census Bureau varies thresholds based upon the size of the family and the ages of the members. The threshold is updated annually using the Consumer Price Index for All Urban Consumers (CPI-U). For a four-person family with two children under the age of 18, the poverty threshold was $20,444 in 2006.

Internationally, income poverty definitions differ remarkably from country to country. In 1990, the World Bank, to assess poverty in developing countries, adopted a measure of US$370 per person per year (in 1985 dollars) as a rule-of-thumb benchmark.

It is commonly referred to as the "dollar per day" poverty line. This indicator has now fallen out of favor in the international development community.

The poverty line, as adopted by the U.S. government, has generated significant controversy. Some have argued that the poverty line does not fully describe the experiences of those in or near poverty, as the proportion of expenditures on housing relative to overall income has changed dramatically over the years. These changes have resulted in a decrease in the amount of disposable income that can be used for other basic needs. Geographical variations in housing costs are also not accounted: The line is the same whether one lives in Manhattan or in rural Arkansas. The line has also been criticized because other costs that decrease disposable income are not accounted (e.g., transportation and child care costs). Changes in family composition and the definition of family over time have also been noted. Others have argued that income definitions do not account for in-kind transfers and governmental support programs that do not fall under the definition of income (e.g., housing allowances and food stamps). Many suggestions have been forwarded and conferences convened that have attempted to create a measure of poverty that more accurately reflects the number of individuals and families that struggle with meeting basic human needs. In 1995, the National Academy of Sciences (NAS) introduced a series of experimental measures and recommendations for assessing poverty in the United States. In short, the participants suggested that the poverty line be based upon a budget for a series of basic needs (food, clothing, and shelter with utilities) and small additional amounts allocated for personal needs). Consumer expenditure data would be assessed to determine the cost of this basic bundle of goods and services and would be updated annually. Geographic variations would be incorporated, along with differences in family composition, for the purpose of determining the threshold. To determine if a family met the threshold for poverty, in-kind and near-cash governmental assistance (e.g., food stamps) would be counted as resources along with the traditional definition of income. These transfers are included because the assistance can be used to purchase goods and services within the basic needs budget. However, expenditures for child care, work-related transportation, and other personal expenses (e.g., health care) would be subtracted from the amount of disposable resources.

To date, these suggestions have not been implemented, though the U.S. Census Bureau reports the National Academy of Sciences alternative figures annually. Though the debate continues, Orshansky's definition has remained the metric of choice, with minor variations, for more than 40 years.

Poverty in the United States

Taken from the Current Population Survey, 2006 Annual Social and Economic Supplement, the U.S. Census Bureau estimates that approximately 37.0 million individuals were in poverty in the United States in 2005. This figure represented 12.6% of the entire population. The highest rate was calculated in the first year of measurement, 1959, when an estimated 22.4% of the population fell below the poverty line. The lowest point that it measured was in 1973, when 11.1% of the population was estimated to be below the poverty line. As recently as 2000, poverty had been as low as 11.3% of the population.

Persons of color represented a disproportionate number of poor individuals in the United States. African Americans and Hispanics had higher rates of impoverishment relative to non-Hispanic Whites. Whereas among Whites, 8.3% were deemed to be below the poverty line, among African Americans, 24.9% were estimated to be below the poverty and Hispanics 21.8%.

Children also represent a disproportionate number of those who fall below the poverty line. For children under the age of 18, 17.6% fell below the poverty line, representing 12.9 million children. For 18- to 64-year-olds the poverty rate was 11.1% and among persons 65 and older, 10.1%.

Poverty and Education Outcomes

Poverty has been linked to greater likelihoods of poor child development from infancy onward and poor education outcomes on a variety of measures. An exhaustive review of the literature by Jeanne Brooks-Gunn and her colleagues of the research surrounding poverty and education/child development outcomes suggests that income poverty has extremely deleterious effects on child development and subsequent education success. In the prenatal and infancy period, low income is associates with a host of poor outcomes, including timeliness of prenatal care, smoking during pregnancy, and low birth weight. Among children

who were born at a low birth weight, low income has been associated with continued poor education outcomes. In the early childhood years intelligence test scores and behavior problems are negatively influenced by income poverty. These poor outcomes continue to affect children throughout their development.

In later years of development, income poverty has been shown to be associated with the likelihood of high school dropout, independent of other factors related to poverty. Scholars have found that the amount of time a family spent below the poverty line was associated with an increased likelihood of dropout. More recent research by Guang Guo and Kathleen Mullan Harris suggests that the impact of poverty on intellectual development is mediated by home environment mechanisms that are directly affected by poverty. Using a structural equation model, Guo and Mullan found that the effect of income on intellectual development was completely mediated by the proposed factors. They concluded that the cognitive stimulation in the home and parenting style were the most significant mediating factors that affected intellectual development in children. The authors suggested that in lieu of income transfers, interventions designed to address these issues in the home may be a viable substitute.

Royce A. Hutson

See also Cultural Diversity; Ethnicity and Race

Further Readings

Bradshaw, T. K. (2006, February). *Theories of poverty and anti-poverty programs in community development* (Rural Poverty Research Center [RPRC] Working Paper No. 06–05). Retrieved May 15, 2007, from http://www.rprconline.org

Brooks-Gunn, J., Duncan, G. J., & Maritato, N. (1997). Poor families, poor outcomes: The well-being of children and youth. In G. J. Duncan & J. Brooks-Gunn (Eds.), *Consequences of growing up poor*. New York: Russell Sage Foundation.

Chambers, R. (2006, December). What is poverty: Who asks? Who answers? *Poverty in Focus*, pp. 3–4.

Fisher, G. M. (1997) *The development of the Orshansky poverty thresholds and their subsequent history as the official U.S. poverty measure* (U.S. Census Bureau Poverty Measurement Working Paper). Retrieved April 24, 2007, from http://www.census.gov/hhes/www/povmeas/papers/orshansky.html

Herrnstein, R. J., & Murray, C. (1994) *The bell curve: Intelligence and class structure in American life.* New York: Free Press.

Iceland, J. (2003) *Poverty in America: A handbook.* Berkeley: University of California Press.

Teachman, J. D., Paasch, K. M., Day, R. D., & Carver, K. P. (1997). Poverty during adolescence and subsequent educational attainment. In J. Brooks-Gunn & G. J. Duncan (Eds.), *The consequences of growing up poor* (pp. 132–189). New York: Russell Sage Foundation.

U.S. Census Bureau. (2006). *Age and sex of all people, family members and unrelated individuals iterated by income-to-poverty ratio and race*. Washington, DC: U.S. Census Bureau. Retrieved May 20, 2007, from http://pubdb3.census.gov/macro/032006/pov/new01_100.htm

U.S Department of Labor, Office of Policy Planning and Research. (1965). *The Negro family: The case for national action*. Retrieved May 21, 2007, from http://www.dol.gov/oasam/programs/history/webid-meynihan.htm

Wilson, W. J. (1987). *The truly disadvantaged: The inner city, the underclass, and public policy*. Chicago: University of Chicago Press.

PRAXIS™

PRAXIS™ is a series of exams and assessments used for determining teacher certification and licensure in 43 states as well as the District of Columbia and Guam. The exams are administered by Educational Testing Service (ETS), a private, nonprofit organization. The PRAXIS Series™ is a trademark of Educational Testing Service. ETS administers the exams on specific dates each year. State education departments determine individually which exams and assessments will be used for their teacher certification process. Each state sets its own passing scores, or qualifying scores. PRAXIS™ raw scores (questions answered correctly) are converted to scaled scores.

PRAXIS I® Exams: Pre-Professional Skills Assessments

The PRAXIS I® (Pre-Professional Skills Test, or PPST®) exams are used by many colleges and universities to evaluate individuals who would like to enter their teacher education programs. The PRAXIS I® PPST® exams are typically taken early in the student's college career.

Each state sets their own qualifying scores, and many educational institutions require students to meet qualifying scores to enter their teacher education program. The PRAXIS I® PPST® exams are available in two formats: paper based and computer based

(offered by appointment through a national network of Prometric™ Testing Centers). The PPST® exams measure basic skills in reading, writing, and mathematics.

The PPST® *reading exam* measures skills in literal comprehension, the ability to clarify a written passage and understand how the material is organized, and the ability to make reasoned, qualitative judgments about the nature and merits of a written passage. What follows is a sample question from the ETS *Tests at a Glance* for the PPST® reading exam:

> Marguerite Duras' achievement as a filmmaker was marked by refusal to become a professional of the cinema, with all that this implies in terms of prestige, influence, financial backing, and even know-how. Although she made many films, she said that she knew very little about the technology of cinema and that she had no reason to learn any more: "I want to remain where I am, on the first grounds of cinema, in the primitive zones."
>
> The passage is primarily concerned with
> (A) condemning critics' failure to appreciate the work of a particular filmmaker
>
> (B) describing the attitude of a particular filmmaker
>
> (C) analyzing the style of a particular filmmaker
>
> (D) criticizing the technical shortcomings of a particular filmmaker
>
> (E) discussing the content of the works of a particular filmmaker

The PPST® *writing exam* assesses the test-taker's ability to use the English language appropriately. The test-taker must be able to identify proper grammar usage in multiple-choice questions as well as compose a written essay. A sample essay topic from the ETS *Tests at a Glance* for the PPST® writing exam is the following:

> Minimum-wage jobs are a ticket to nowhere. They are boring and repetitive and teaching employees little or nothing of value. Minimum-wage employers take advantage of people because they know they need a job.
>
> Discuss the extent to which you agree or disagree with this opinion. Support your views with specific reasons and examples from your own experience, observations, or readings.

The PPST® *mathematics exam* measures mathematical skills that an educated adult or teacher might need. Test-takers must be able to solve problems and reason in a quantitative process. The ability to interpret graphs and determine ratios, percentages, and measurements are components of the exam. Test-takers must also be able to practice informal geometry and use other formal methods of mathematical reasoning. A sample question from the ETS *Tests at a Glance* for the PPST® mathematics exam is this:

> Which of the following fractions is least?
> (A) 11/10
>
> (B) 99/100
>
> (C) 25/24
>
> (D) 3/2
>
> (E) 501/500

PRAXIS II® Exams: Subject Assessments

The PRAXIS II® exams assess specific content knowledge that educators will teach in the classroom as well as teaching skills. States using the PRAXIS II® exams require candidates who are seeking initial teacher certification to meet qualifying scores for licensure. PRAXIS II® exams are Subject Assessments (pertaining to general and specific content knowledge and skills), Principles of Learning and Teaching Tests (pertaining to general pedagogical knowledge) and Teaching Foundations Tests (pedagogy in five areas: multisubject [elementary], English, language arts, mathematics, science, and social science). In many states, colleges and universities encourage teacher candidates to take and pass their appropriate PRAXIS II® exams before their required student teaching. A sample question from the ETS *Tests at a Glance* for the Social Studies Content Knowledge PRAXIS II® exam follows:

> Which of the following is an example of a concurrent power:
> A. The printing and coining of money
>
> B. The power to declare war
>
> C. The process of naturalization
>
> D. The levying of taxes

PRAXIS III® Assessments

PRAXIS III® assessments are classroom performance assessments that are actually completed within a classroom setting to assess the skills of a beginning teacher in some states. The PRAXIS III® system uses a three-pronged method to assess the beginning teacher's evidence of teaching practice that includes direct observation in the classroom, review of documentation prepared by the teacher, and interviews with trained assessors. PRAXIS III® helps teachers identify pedagogical areas in which they may benefit from additional development. The PRAXIS III® Classroom Performance Assessments contains 19 assessment criteria in four interrelated domains. These domains embrace the teaching and learning experiences of the beginning teacher in curriculum planning, creation of learning environments, instructional methods, and teacher professionalism.

Tracy Hack

See also Assessment; Evaluation; Testing

Further Readings

Praxis® I overview. Retrieved August 27, 2007, from http://www.ets.org

Rikard, L., & Norden, L. (2006). Winning strategies for passing Praxis I and II. *Journal of Physical Education, Recreation & Dance, 77*(3), 44–47.

Watters, J. J., & Diezmann, C. M. (2007). Multimedia resources to bridge the Praxis gap: Modeling practice in elementary science education. *Journal of Science Teacher Education*, *18*(3), 349–375.

PRECISION TEACHING

Precision teaching is a method for learning and practicing skills and strategies until they are fluent or automatic. It was developed in the mid-1960s by Ogden Lindsley, who had been a student of B. F. Skinner at Harvard University in the 1950s. Precision teachers assert that these procedures ensure that students retain the skills they are taught, can perform them for extended periods, and can easily apply them, both to new learning requirements and in the course of their daily lives. This entry describes the evolution of precision teaching and its accompanying standard celeration chart.

History

Lindsley began by developing a graph, which he initially called the standard behavior chart. His vision was that teachers could be scientists, plotting each student's performance data on the chart, watching the trends in each student's performance as they learned, and prescribing interventions to improve their growth. By always using the same chart, teachers could also communicate the results of their efforts with each other in a standard way instead of inventing new graphs for each communication. After a few years, Lindsley developed an orderly system for using the chart, which he called *precision teaching*.

Precision Teaching Method

The method prescribes five steps. First, teachers define a learning objective, which he called a *pinpoint* because it precisely pinpoints a skill, fact, concept, or principle to learn. Second, they arrange materials and procedures for learning and practicing the pinpoint. Precision teachers who are teaching a group of students usually provide instruction on a skill and then give students worksheets, flashcards, or some other arrangement to practice that skill. Precision teachers who are tutoring individual students often begin immediately with timed practice, guiding the student's performance with "tips and quips" and cheerleading. Third, teachers and students time the student's performance and count its frequency. The *timing* may be as short as 10 seconds or as long as 15 minutes or more. Students increase their frequencies by repeatedly practicing a skill in a series of short timings, often 1 minute in length, over successive days. Students in classroom settings often work in pairs, with one student monitoring the other's performance and giving feedback as they take turns practicing. A fluency aim is set for each skill, specifying the

frequency of tasks that should be completed in the timing period. Fourth, students and teachers chart the student's performance on the standard celeration chart. They then compare their frequency to the fluency aim and keep practicing until they reach it. Depending on the skill, they could reach fluency in 1 day, several days, weeks, or even months. Fifth, students and teachers review trends of performance on the chart and make decisions about possible interventions to improve performance. Some precision teachers set minimum celeration aims to help ensure that students achieve fluency aims in a timely manner. Peers coach and cheerlead each other to do better on each successive timing. Lindsley developed a motto for the method: "Pinpoint, record and chart, change, try try again."

Features

Performance Measurement

Several features of precision teaching are noteworthy. First, the focus is on measuring performance. Lindsley was confident that if teachers keep precise track of how their students are learning, they will figure out what to do to improve performance. Thus, precision teaching is not directly a teaching method; rather, teaching methods derive from the precise measurement of learning.

Frequency Measurement

Second, Lindsley prescribed measuring each student's frequency of performance in a period of time (count per minute). Frequency measurement is much closer to direct observation of behavior than are derivations such as percentage correct. Frequency is a sensitive measure of behavior change. In one research project, three groups of subjects (children and adults with developmental disabilities, public school elementary students, and teachers and psychologists) engaged in simple tasks such as copying numerals 0 to 9, naming pictures, counting by rote, and putting tiles into a can. Performance was measured for 1 minute, and the median frequencies for each group were calculated. There were no differences in the percentages of correctly completed tasks among the three groups; members of each group performed the tasks at 100% correct. However, when performance frequency was plotted, there were large differences between each

group: students with developmental disabilities performed between 5 and 20 per minute, public school students performed 20 to 40 per minute, and teachers and psychologists performed more than 100 per minute. The differences one would intuit among the groups were masked by the grosser percent correct calculations but revealed by measuring frequency.

Standard Celeration Chart

A third noteworthy feature of precision teaching is its use of a special graph or chart, currently called the *standard celeration chart*. The standard celeration chart is a tool that embodies the characteristics of measurement typical of the natural sciences. In natural science, units such as distance (e.g., centimeters), mass (e.g., grams), and time (e.g., minutes) have three properties. First, they are standard, having the same meaning wherever and whenever they are used. Second, they are absolute; that is, their value does not change from one use to the next. Third, they are universal, applying to every instance of the unit as it occurs in nature. Frequency (count per minute) and time (calendar days) are standard, absolute, and universal aspects of behavior. Lindsley created a chart with frequency on the y axis and calendar days on the x axis. Users convert and plot all frequencies as "count per minute," regardless of the actual time allotted in a timing—-another standard feature of the chart. All human behavior frequencies can be plotted on a standard celeration chart, from .001 per minute (approximately once a day) to 1,000 per minute.

When frequencies are charted on successive days, a line with a slope is created. The slope is a measure of celeration, that is, (count/time)/time, a speed measure divided by time. *Celeration* is the root word of *acceleration* and *deceleration*, terms from algebra and physics. The chart creates a picture of the change in frequency over time. Precision teachers use the chart to quickly see celerations in frequency and make interventions to alter them if necessary. The celeration picture on the chart is standardized. For example, a celeration line that forms a diagonal slope shows a doubling in celeration across time. The doubling celeration can occur in any frequency range. Because the picture is standard, teachers can look at any chart and tell at a glance how learning is improving and whether they need to make interventions to correct the learning trend, that is, accelerate correct performance or decelerate incorrect performance.
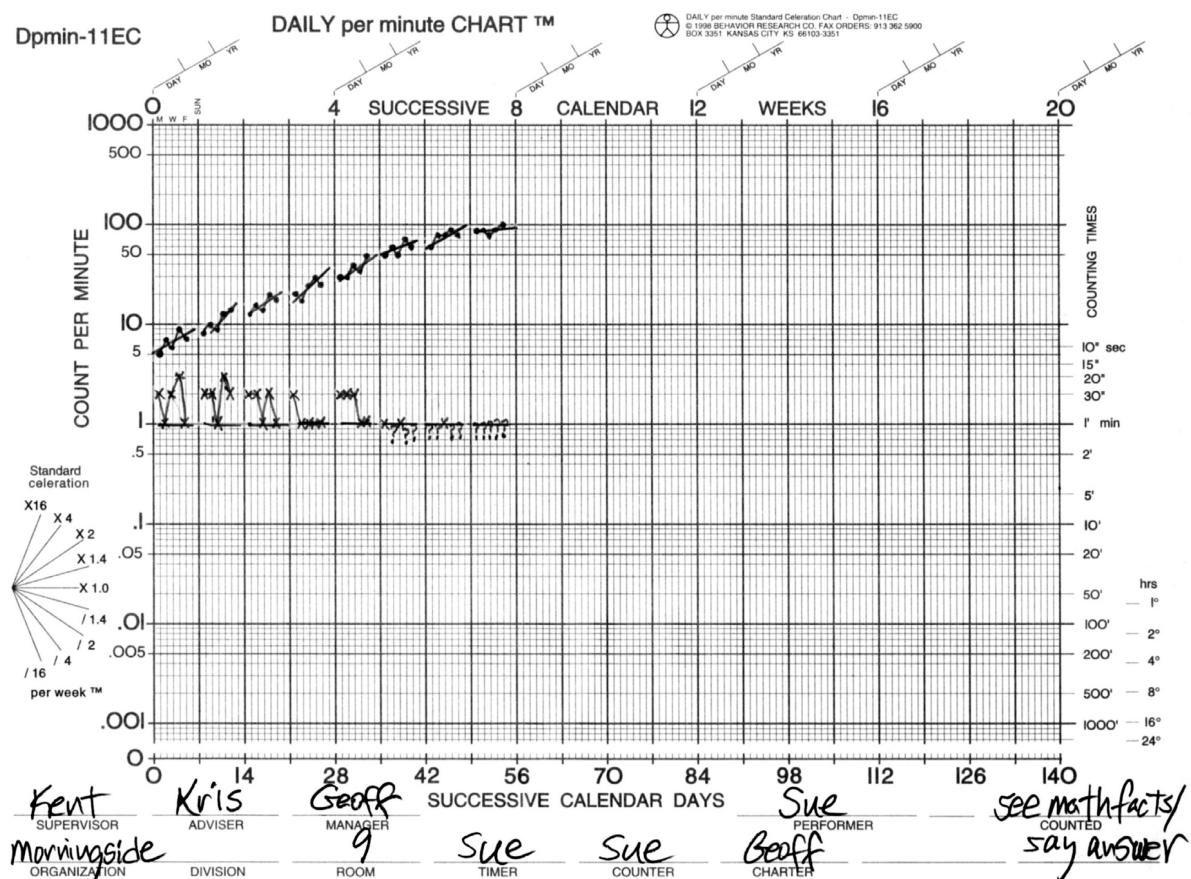
Figure 1   Standard Celeration Chart

Note: This chart shows a student's frequency of performance over successive days, with celeration lines indicated for each week.

Lindsley created a family of charts: "daily" (count per minute), "weekly" (count per week), "median count per minute per week," "monthly" (count per month), "monthly summary," "yearly" (count per year), and "timing charts" on which students and teachers graph frequencies for each timing completed.

Standard celeration charts are used by a wide variety of people. They can be used in a one-on-one arrangement or with a group of people, each of whom keeps his or her own chart. Teachers record performance and decide whether to change curricula or add interventions. Marriage counselors help couples track progress in their relationships and decide whether to make life changes. Sales people measure the effectiveness of different campaigns. Administrators, CEOs, planners, policymakers, and politicians detect historical and economic trends

and project future conditions. Chart users can compare data across individuals, schools, industries, laboratories, and disciplines, without needing any other charts or graphs.

In the first decade of precision teaching, teachers measured and charted the full duration of the tasks they assigned to students. For example, if students were assigned to read a passage in a textbook, teachers would record the frequency of words read in the time it took to read the passage. In the 1970s Eric Haughton experimented with using brief, 1-minute timings for classroom learning of basic skills such as reading and arithmetic and compared them to measurements of the full duration of the performance. Would tracking performance in brief timing periods be sufficient to guide the teacher's teaching, decision

making, and interventions for student growth? Indeed, for many pinpoints, 1-minute timings were sufficient and provided other benefits as well. Short timings could be repeated and charted, and a measure of growth (or celebration) incorporating more data points could be determined.

Learning Aims

Haughton also experimented with setting performance standards—aims or goals that students should achieve during learning and practice. Achieving these aims indicated fluency of the performance. A formal definition of *fluency* evolved that related to multiple learning outcomes. It was Haughton's conviction that students should practice a performance until they achieved certain frequency aims, because frequency building produces fluency. Aims were set as a function of whether reaching them predicted that the skill would be (a) retained, (b) enduring (i.e., the students could engage in the skill for longer periods of time without fatigue), and (c) easily applied to new and more complex learning. Haughton stated an acronym as a challenge to guide teachers: Teach your curriculum to REAPS (Retention Endurance Application Performance Standards). He encouraged teachers to set goals for these and other learning outcomes. His challenge to teachers was to teach to multiple, high-performance standards. In recent years, others have added to the acronym, for example, practice skills until they are stable (i.e., not subject to distraction) and practice skills until they easily combine with other skills as necessary to figure out how to engage in a more complex skill that has not yet been taught (RESAAPS: Retention, Endurance, Stability, Application, Adduction Performance Standards).

Thus, precision teachers set frequency aims for fluency building. For example, students of precision teachers practice until they can complete math facts at 80 to 100 per minute with no errors, solve 12 to 15 math word problems in 5 minutes with no errors, add and subtract fractions at the rate of 80 to 100 steps in 5 minutes with no errors, orally read 180 to 220 words per minute with two or fewer errors, find the main idea in 8 to 10 short passages in 5 minutes with no errors, and write persuasive paragraphs at 20 words per minute.

Precision teachers may also set celeration aims for student performance. A celeration aim indicates the amount of time it takes to reach a frequency aim. For example, a teacher may set a celeration aim at "times 2" for learning math facts, requiring a doubling in performance per week. Teachers and students draw a minimum rate-of-improvement line on their charts to indicate times-2 growth. The slope of the celeration line indicates how quickly or timely the student is to achieve fluency. As students practice, they plot their own improvements and compare their progress to the minimum rate lines. Their comparisons tell them whether they are making sufficient progress to reach fluency in a timely manner or whether they need to ask the teacher or another student for help.

Headsprout Early Reading is an Internet-driven, cartoon-based, beginning reading program that automates the setting of frequency and celeration aims. Student performance frequency is directly tracked online as the student proceeds through the program. The program progresses as a function of when the learner meets or beats fluency and celeration aims.

Setting frequency and celeration aims has implications for the arrangement of practice. Precision teachers make sure that students have enough exercises to complete in the time they are allotted. For example, to reach an aim of 80 to 100 math facts per minute with no errors, students need more than 100 problems at their disposal. To provide students with fewer than 100 problems places what precision teachers call a *ceiling* on the frequency a student can achieve. For example, if the teacher gave the student only 30 math facts to complete in 1 minute, the highest frequency they could chart would be 30. Students learn that they cannot possibly complete all of the exercises a precision teacher gives them for completing a timing.

With precision teaching, teachers can arrange conditions to teach important goal setting, self-monitoring, self-management, organizational, and cooperative learning skills. For example, students may learn self-management and self-determination through freedom to take their own performance breaks and still meet their expected goals, skip lessons when they can demonstrate mastery, move through the curriculum at their own pace, select their own arrangement of tasks to accomplish in a class period, and choose their own free time activities.

Kent Johnson

See also Applied Behavior Analysis; Effective Teaching, Characteristics of; Operant Conditioning; Special Education; Teaching Strategies

Further Readings

Binder, C. V. (1996). Behavioral fluency: Evolution of a new paradigm. *The Behavior Analyst*, *19*, 163–197.

Johnson, K. R., & Layng, T. V. J. (1992). Breaking the structuralist barrier: Literacy and numeracy with fluency. *American Psychologist*, *47*, 1475–1490.

Johnson, K. R., & Street, E. M. (2004). *The Morningside model of generative instruction: What it means to leave no child behind.* Concord, MA: Cambridge Center for Behavioral Studies.

Lindsley, O. R. (1972). From Skinner to precision teaching: The child knows best. In J. B. Jordan & L. S. Robbins (Eds.), *Let's try doing something else kind of thing* (pp. 1–11). Arlington, VA: Council for Exceptional Children.

Lindsley, O. R. (1990). Precision teaching: By teachers for children. *Teaching Exceptional Children*, *22*(3), 10–15.

Pennypacker, H. S., Gutierrez, A., & Lindsley, O. R. (2003). *Handbook of the standard celeration chart* (deluxe ed.). Concord, MA: Cambridge Center for Behavioral Studies.

PREMACK PRINCIPLE

B. F. Skinner developed the operant conditioning learning framework to explain how organisms work to receive rewards and avoid punishment. Skinner also noted, via the results of animal experiments, that behaviors could be changed in relation to the consequences paired with the behaviors. Specifically, behaviors are responded to in one of two ways: reinforcement or punishment. Unlike punishment, which aims to decrease the frequency of a behavior, reinforcement seeks to increase the desired behavior's frequency. Positive reinforcement occurs when a pleasant stimulus is added following the occurrence of a behavior, such as giving a treat to a dog every time it successfully follows a command to sit. Negative reinforcement occurs when an aversive stimulus is removed following the occurrence of a behavior, such as when pressing a button (behavior) causes loud noises (aversive) to be taken away. Despite the widespread applicability of operant conditioning, there are limitations of the framework. Particularly, operant conditioning does not detail *why* an organism carries out a low-frequency behavior and does not delineate what reinforcers are.

David Premack acknowledged this limitation and conceptualized a way to explain how to reinforce a low-frequency behavior by using an existing high-frequency behavior. He stated, "Any response A will reinforce any other response B, if and only if the independent rate of A is greater than that of B" (p. 31). Premack termed this notion the *Premack Principle* (also known as the *differential probability principle*). He explained that the principle operates based on a reinforcement hierarchy, which is a list of actions that are ordered from the highest to the lowest probability behaviors. To further illustrate the principle it is useful to consider the work of W. J. Johnson. Johnson used the Premack Principle to encourage a young man with depression to practice positive self-talk. He told the young man to trigger a positive thought (low-frequency behavior) from reading a statement on an index card every time he urinated (high-frequency behavior). After the young man continued this exercise for several weeks, he began to think of positive statements without the index cards before urinating. Soon after that, he was engaging in positive self-talk on a regular basis and not solely in the context of the high-frequency behavior. In this example, positive self-talk could not have been learned and reinforced if the experimenter had paired it with behavior that occurred at a lower frequency than the current rate of self-talk.

Premack further described the reinforcement hierarchy by stating that if A, B, and C are any three responses, with A having the highest frequency followed by B and then C, then A will reinforce both B and C, B will reinforce C but not A, and C cannot reinforce either A or B. The Premack Principle also elaborates on contingency (i.e., dependence). Organisms are taught through trial and error that relationships exist between behaviors and the environment. Specifically, certain behaviors produce particular changes in the organism's environment, so that changes in the environment are contingent upon the identified behaviors. In reference to the Johnson example, if urinating is contingent upon thinking of positive thoughts, then positive thoughts will occur. In his own experiment, Premack allowed a group of children access to a pinball machine and candy. He observed which behavior occurred with higher frequency: manipulation of the machine or eating of the candy. Nearly half of the children were more interested in eating candy than manipulating the machine, whereas the remainder of the children displayed the opposite pattern. For the children that preferred playing with the pinball machine, experimenters made it so the children would have to eat the candy to be able to manipulate the pinball machine. Conversely, for

the children who preferred eating candy, experimenters told them they had to play the pinball machine if they wanted to eat candy. Results showed that when the high-frequency behavior of the child was contingent upon the low-frequency behavior, the child would engage in the low-frequency behavior. This experiment supported the Premack Principle in that it demonstrated that organisms could be influenced to carry out a behavior if the desired behavior is dependent upon the less-desired behavior.

The utility of the Premack Principle is undeniable. That is, Premack incorporated components of Skinner's operant conditioning framework and went further by detailing the process of reinforcing low-frequency behaviors and identifying an organism's reinforcers. Moreover, the principle highlights the ability to manipulate or motivate the behaviors of organisms by utilizing high-frequency behaviors already in existence. The principle can be enacted in a variety of settings (e.g., home, school) and may be modified dependent on which high-frequency behaviors typically occur in particular settings.

Rachael Elizabeth Kroening and Kerri Lynn Kim

See also Applied Behavior Analysis; Behavior Modification; Classical Conditioning

Further Readings

Martin, G., & Pear, J. (1999). *Behavior modification: What it is and how to do it* (6th ed.). Upper Saddle River, NJ: Prentice Hall.

Premack, D. (1959). Toward empirical behavior laws. *Psychological Review*, 66, 219–233.

Private Speech

Private speech (also known as *private dialogue, self-talk, collective monologue, noncommunicative speech,* or *egocentric speech*) is speech that is directed to the self and not to others. Private speech serves as a tool for thinking that facilitates problem solving and is also a tool that mediates people's self-regulatory capacity. John Flavell coined the term *private speech* in 1964 as a means of distinguishing Lev Vygotsky's understanding of this phenomenon from the interpretation given to us earlier by Jean Piaget. The earliest documentation of egocentric speech surfaced in the

1920s by Piaget. He observed young children playing together while emitting verbal monologues without any apparent concern for the listener. The children appeared to be unable to take the perspective of the listener into account and instead were verbalizing their thoughts on two very separate topics. Jean Piaget reasoned that these collective monologues were evidence of children's cognitive and social limitations. These limitations caused young children to converse in an egocentric manner. Piaget reasoned that such speech was evidence of immature social or communicative speech.

Soon after, Vygotsky became aware of Piaget's observations and offered another interpretation for this so-called egocentric speech. Vygotsky believed that private speech was a qualitatively distinct type of speech directed to the self and provided specific functions to individuals rather than a communicative function with others. This self-talk was a mediational means by which young children are able to collaborate with themselves to solve tasks, motivate themselves toward attaining goals, and regulate their behavior. Vygotsky argued that egocentric speech was actually a tool for thinking and represented the point in development when two functionally separate systems, developing parallel with each other, converge, thus allowing our thought system to become verbal and our speech system to become rational. Vygotsky also believed that this new qualitatively distinct form of speech was essentially overt inner speech. Private speech is the precursor to covert verbal thought and, with time, goes "underground" to form inner speech.

It is important to understand the role that this developmental phenomenon plays in children's development of self-regulation. Vygotsky suggested that private speech was the means by which higher psychological processes like voluntary control and voluntary attention develop in young children. There exists a significant body of empirical literature that supports Vygotsky's view of private speech as a tool of thought, and there is an emerging body of literature that also supports Vygotsky's view that private speech is the basis of people's motivations.

Contemporary research has shown that private speech plays a critical role in the transition from external social regulation to internal, autonomous self-regulation. Private speech emerges soon after children are able to engage in social, communicative speech. Once children are able to communicate, they begin to communicate with themselves as a means of

planning, guiding, and orchestrating their cognitions and actions. Research has shown that children begin to use private speech at approximately 2 years of age and increase their use to about 5 to 6 years of age. After the age of 6 or 7, private speech diminishes as children are able to engage in dialogue with themselves covertly. Although it was once thought that private speech would disappear after the age of 7 or 8, current research on older children and adults has shown that it emerges relative to task demand and individual's need to self-organize and self-direct thoughts, motivations, and actions.

Private speech also has been investigated in toddlers, adults, and the elderly. Private speech has been shown to serve various self-regulatory functions in a number of task situations and in a number of contexts. It is dynamically related to children's current and future task performance and is produced relative to task demand. Private speech is also dynamically related to social situations and, in particular, to the type and quality of guidance children receive while working with more competent peers or adults in a scaffolded manner. Private speech is also produced relative to one's perceived self-efficacy and one's inner resources for motivating a course of action.

Many researchers in their study of private speech have developed category systems. Many category systems differentiate the content of private speech utterances, whereas other systems differentiate between hypothesized functions of private speech. Other category systems have differentiated between the various forms or physical properties of private speech utterances, whereas others have examined the variety of environmental contexts in which private speech is emitted.

Private speech has been studied in many cultures and by numerous international scholars in the private speech field. Qualitative and quantitative methodologies have been used in the design of private speech studies. Direct observational methods have been used as well as survey/questionnaire methods to gather information about children's and adult's use of private speech. Various task situations have been developed to study private speech and various social situations also have been used. Both laboratory and naturalistic settings have been used and comparisons of children's indoor and outdoor learning environments have been conducted. Researchers have learned that private speech is related to language development, memory processing, bilingualism, academic motivation, self-awareness,

creativity, and problem solving in the classroom. Current private speech research programs have expanded their attention on private speech to examine its relationship to children with special needs (e.g., those with autism spectrum disorders or attention deficit hyperactivity disorder), whereas others have incorporated recent advances in neurobiology and cognitive neuroscience, such as functional magnetic resonance imaging and positron emission tomography. Current private speech investigators are also studying how private speech mediates the development of executive functions, action control, and academic motivation.

David J. Atencio

See also Language Disorders; Social Development

Further Readings

Berk, L. E. (1994, November). Why children talk to themselves. *Scientific American*, 78–83.

Diaz, R. M., & Berk, L. E. (Eds.). (1992). *Private speech: From social interaction to self-regulation.* Hillsdale, NJ: Lawrence Erlbaum.

Winsler, A., Fernyhough, C., & Montero, I. (2007). *Private speech, executive functioning, and the development of verbal self-regulation.* New York: Cambridge University Press.

PSYCHOANALYTIC THEORY

Psychoanalytic theory began with the work of Sigmund Freud, often considered the founding father of psychology, in the late 19th century. Psychoanalytic theory is broad in scope and offers unique, controversial insights into how the human mind works. Freud's work has been the catalyst for the development of other psychodynamic theories and other theories that developed out of resistance and opposition to psychoanalytic notions. Psychodynamic theories and ideas have permeated our society. This entry (a) provides an overview of psychoanalytic theory, (b) describes the core concepts of its structural model of personality, (c) describes the purported developmental stages associated with psychoanalytic theory, (d) presents a critique of the theory, and (e) provides a brief overview of other theories that have developed out of psychoanalytic theory.

Overview

According to Freud, there are three components to the human mind: the conscious, which represents our present awareness, including our thoughts, feelings, and perceptions; the preconscious, which represents things we are not presently aware of but can become aware of with little effort; and the unconscious mind, which represents things we are unaware of but which still affect our functioning. The unconscious mind includes both forces that originate there (our basic instincts and drives) as well as unacceptable thoughts and memories that may have been pushed back because of their threatening nature.

The core of psychoanalytic theory is that all behavior, thoughts, and emotions are influenced by unconscious processes and internal forces. These internal forces are considered to be dynamic in nature and thus always changing. The ways in which the internal forces interact with one another gives rise to behavior, thoughts, and emotions. When the forces are in conflict, abnormal behavior arises. Psychoanalytic theory has a deterministic assumption, meaning that no symptom or behavior is accidental, but rather it is all determined. In particular, it is determined by past experiences, particularly those during early childhood.

Structural Model of Personality

Freud believed that there are two basic forces that motivate human behavior: the sexual drive (referred to as the *libido*) and the aggressive drive. The energy created from these drives continually seeks to be released but can be channeled by psychological systems. The psychological systems that help regulate the drives are the id, the ego, and the superego.

The id is the innate system from which the libido and aggressive drives emerge. The id is located in the unconscious mind and operates from the pleasure principle, which seeks immediate gratification of sensual needs and drives. Freud viewed infants as operating on nearly purely the id. For example, an infant who is hungry will seek his or her mother to breastfed or scream and cry if the mother is not available. When direct action cannot be taken, fantasies or memories are created to meet the id's needs. In the case of the infant, the infant may imagine the sight of his or her mother's breast. This process is referred to as *primary process thinking*, or wish fulfillment. The id is loosely associated in organization, holds

a distorted cognitive representation of the world, and is best seen in dreams.

As children develop, they begin to recognize that they cannot have immediate wish fulfillment without potentially negative consequences. Over time they develop the ego, a part of the psyche that seeks to fulfill wishes in more appropriate and socially acceptable ways. The ego is the center of consciousness and follows the reality principle, which is the drive to satisfy needs within the limits of society's rules, rather than the pleasure principle. The ego attempts to regulate conflict between the instinctual drives of the id and the demands of the external world and, in doing so, seeks maximal gratification of instincts while maintaining relationships with the outside world. The ego uses *secondary process thinking*, or rational deliberation, as its primary mode of operation. For example, a preschooler who wishes to continue breastfeeding may realize this is no longer allowed or considered acceptable and thus may satisfy herself with snuggling with her mother instead.

Later in childhood, the superego develops out of the ego. The superego represents the ethical and moral attitudes of a child's parents and society more broadly. It is purported to develop out of the internalization of the moral and ethical standards of society, and as such, it holds the rules and regulations to be followed. We internalize these moral standards because following them feels good and reduces anxiety. The superego is often the source of feelings of guilt and instructs humans on all of the things they *should* do; thus, it is typically thought of as the heart of our conscience.

The three components—the id, ego, and superego—are frequently in conflict, with competing wishes and demands on the person. The interactions among them typically occur in the unconscious part of the mind, where we are completely unaware. Wishes, desires, needs, and memories occasionally work their way into the preconscious from the unconscious mind, but rarely do they enter the conscious mind. Our unconscious and preconscious work to protect our conscious from the wishes, needs, and memories that represent our basic instincts and drives because these are often unacceptable to the individual person or society at large.

Defense Mechanisms

When unconscious wishes seep into the conscious mind, it can be distressing to individuals. As a way of

coping, individuals develop defense mechanisms to help disguise or transform the unacceptable, unconscious desires. All people use defense mechanisms on occasion to protect themselves from their unconscious desires. Freud argued that the particular pattern of defense mechanisms a person uses helps shape his or her personality. When maladaptive mechanisms are used or when an individual's behavior becomes governed by the defense mechanism, then it can lead to abnormal behavior.

The defense mechanisms described by Freud are as follows:

1. *Repression*—not allowing painful or dangerous thoughts to become conscious (e.g., a person who witnessed a murder may completely repress the memory and have no recollections of the incident)

2. *Denial*—refusing to acknowledge the existence of an external source of anxiety (e.g., a parent insisting her fatally ill child will be all right)

3. *Projection*—attributing one's own unacceptable impulses, motives, or desires to others (e.g., a husband who is sexually attracted to a friend accuses his wife of being unfaithful to him)

4. *Rationalization*—creating socially acceptable reasons for an action that actually reflects unattractive motives (e.g., a student who did not get his job of choice rationalizes that he did not like the city the other job was in anyway)

5. *Reaction formation*—adopting a behavior that is the exact opposite of the impulses one is afraid to acknowledge (e.g., a man who is afraid of being too feminine joins several ''tough, manly'' sports teams, such as football and hockey to prove his masculinity)

6. *Displacement*—displacing hostility away from a dangerous object and onto a safer substitute (e.g., a wife is angry at her boss but cannot safely express her anger to him and so she goes home and yells at her family)

7. *Regression*—retreating from an upsetting conflict to an earlier developmental stage to prevent anxiety and assuage current needs (e.g., a previously toilet-trained preschooler begins to have accidents after his new baby brother is born)

8. *Identification*—adopting the ideas and values of someone in a superior position (e.g., prisoners adopting the attitudes of their captors)

9. *Sublimation*—expressing sexual or aggressive energy in ways that are acceptable to society (e.g., a man with strong aggressive impulses plays professional football)

10. *Intellectualization*—adopting cold, clinical, distanced perspectives on issues that in reality create strong, unpleasant feelings (e.g., a husband going through a divorce is extremely emotionally distressed and daily discusses the technical logistics of court proceedings with colleagues, with an emotional detachment)

Although rigid repeated use of defense mechanisms can be problematic, not all of these defenses are unhealthy coping strategies. Within sublimation are love, work, altruism, and even humor—all of which involve rechanneling raw sexual and aggressive impulses. According to Freud, love is an especially powerful form of sublimation because it allows people to achieve sexual gratification in a socially acceptable context. In fact, all forms of sublimation were viewed as constructive mature defense mechanisms by Freud, and research by George Vaillant has demonstrated that people who engage in such mechanisms are physically healthier, materially more successful, and more resistant to mental illness than people who do not display such defenses.

Psychosexual Stages of Personality Development

Psychoanalytic theory contends that a child's early childhood relationships, particularly those with his or her caregivers, are important influences of personality development. Freud claimed that as children develop, they go through a universal series of psychosexual stages. Each stage of development has psychological conflicts to be addressed by the id, ego, and superego, and each stage focuses on a different sexually excitable zone of the body. The psychological issues and conflict within each stage must be successfully negotiated for the child to become a psychologically healthy adult.

The way a child learns to fulfill the sexual desires associated with each stage becomes an important component of his or her personality, and the caregiver's responses to the child's attempts to satisfy basic needs and wishes can greatly influence whether a given stage is negotiated successfully. If a child does not successfully adjust, then fixation occurs, wherein the child can become trapped at an earlier stage.

The first stage is the oral stage, which occurs during infancy to approximately 18 months of age. In this stage libidinal impulses are satisfied through stimulation of the mouth area, usually through feeding or sucking. If the caregiver is not adequately available, the child can develop a deep sense of mistrust and fear of abandonment. Fixation at this stage results in excessive dependence on others and habits related to the mouth, such as smoking and excessive eating and drinking.

Next is the anal stage, which lasts from age 18 months to approximately 3 years. The focus of gratification in this stage is the anus, with children expressing great interest in passing and retaining feces. If parents are too harsh or critical with toilet training, the child may become fixated at this stage and have traits such as being stubborn, overcontrolling, stingy, and too focused on orderliness.

Following the anal stage is the phallic or Oedipal stage, which occurs from age 3 to 6 years. In this stage, the focus of pleasure is the genitals. The phallic stage is where the most important sexual conflicts occur: the Oedipus complex for boys and the Electra complex for girls. Boys must work through the Oedipus complex, wherein they are purported to fall in love with their mothers and fear their fathers will retaliate by castrating them. The fear of castration is quite strong for boys and is the source of motivation for resolving this conflict. Resolution of this results in a strong superego and identification with their father's values. The Electra complex is the equivalent for girls. Girls' movement through this stage involves the experience of falling in love with their fathers and fearing retaliation by their mothers. However, Freud argued that there is no castration anxiety for girls because they were already castrated. As a result, girls do not have as strong of a motivation for developing their superegos and thus are argued to not be as moral. Freud argued they are instead motivated more by emotions than morals and by penis envy. According to Freud, unsuccessful resolution of the phallic stage results in not adopting appropriate gender roles or a heterosexual orientation and excessive seductiveness in relationships.

Next is the latency stage, which covers middle childhood, from age 6 to 12. During this stage, libidinal drives are more at rest and there is a tendency to avoid the opposite sex. The focus is now more on developing skills and interests. Last, is the genital stage, which begins around age 12 at puberty. If children have successfully resolved their previous stages, their sexual interests turn to heterosexual relationships. During this stage they pursue and develop romantic relationships and learn to negotiate romantic and sexual encounters with the opposite sex.

Beyond Freud

Critiques of Psychoanalytic Theory

Freud's psychoanalytic theory has been the object of numerous critiques over time. In particular, his theory is most often criticized for its biases regarding gender and sexual orientation. Many objections to his work stem from his propositions that girls can never be as moral as boys given that they have never had the strong castration anxiety as motivation. Additionally he is criticized for the idea that girls and women are motivated out of the envy and desire of having a penis rather than by morals and intellectual processes. Other objections focus on his assumptions that homosexuality is abnormal and a result of a fixation caused by early conflict. The idea that heterosexual relationships are the only possible mature outcomes of successful stage resolution is objectionable to many. A third main criticism to his theory is the strong focus on sexuality as the driving force behind development and behavior. Freud developed his theory during the Victorian era with its strong suppression of sexuality, and so it is not a coincidence that his work is so strongly influenced by the context in which he worked. Given the changes in society, however, the theory may not be as applicable. Another criticism centers on notions of determinism. Many people object to the idea that all behavior is predetermined by past behavior, as it does not allow for simple mistakes, is too rigid in assuming people will always act based on what they learned early in childhood, and does not allow for the possibility of change. Finally, a crucial problem with the theory is the difficulty in scientifically testing its fundamental assumptions.

A New Generation of Psychodynamic Theorists

Criticisms aside, many elements of Freud's work can be seen in more modern psychodynamic theories such as object relations, which argues that our early relationships create representations of ourselves and others which influence all of our subsequent relationships. Elements of psychoanalytic theory can also be seen in Carl Jung's theory of the collective unconscious,

which consists of memory traces of past generations rather than just those based on early childhood. Alfred Adler placed less emphasis on the sexual and aggressive drives and viewed the self as serving a more meaningful purpose, such as seeking to fulfill ourselves to fullest extent possible. Others found different areas of emphasis from Freud, such as Karen Horney, who believed that basic anxiety stems from social rather than biological sources, and Harry Stack Sullivan, who focused on the issue of the personality being inextricably intertwined with the social context in which it operates.

Sigmund Freud's psychoanalytic theory was the first comprehensive theory of the mind and personality. It served as a framework for understanding normal and abnormal behavior, provided techniques for addressing abnormality, and spawned a wide range of other theories, including those within the psychodynamic umbrella but which have a slightly different emphasis, as well as those that were developed out of complete objection to his ideas, such as behaviorist and cognitive-behavioral theories. In spite of the many criticisms of this theory, it holds an important place in psychological history and has shaped much of psychology and psychiatry in the past century.

Tina D. Du Rocher Schudlich

See also Attachment Disorder; Erikson's Theory of Psychosocial Development; Individual Differences; Kohlberg's Stages of Moral Development

Further Readings

Burger, J. M. (1997). *Personality* (4th ed.). Pacific Grove, CA: Brooks/Cole.

Demorest A. (2004). *Psychology's grand theorists: How personal experiences shaped professional ideas.* Mahwah, NJ: Lawrence Erlbaum.

Freud, S. (1953–1964). *The standard edition of the complete psychological works of Sigmund Freud* (J. Strachey, Ed., with A. Freud; 24 vols.). London: W. W. Norton.

Freud, S. (1965). *The interpretation of dreams* (J. Strachey, Ed. & Trans.). New York: Basic Books.

McWilliams, N., & Weinberger, J. (2003). Psychodynamic psychotherapy. In G. Stricker & T. A. Widiger (Eds.), *Handbook of psychology: Clinical psychology* (Vol. 8, pp. 253–277). New York: Wiley.

Vaillant, G. E. (2000). Adaptive mental mechanisms—their role in a positive psychology. *American Psychologist*, *55*, 89–98.

Psychosocial Development

Psychosocial development can be defined as the process by which an individual attempts to become part of a society while maintaining his or her individuality. This process does not occur with the individual being the sole actor but rather occurs through the act of socialization. *Socialization* is how an individual develops the values, behaviors, and knowledge about the societies in which that individual lives, and because a person lives in multiple simultaneous societies and subsocieties (communities, schools, homes, peer groups, etc.), it can be a complex task to integrate all of them.

Two types of psychosocial development exist. The first, social roles, helps the individual understand how to act appropriately in his or her social world, as social roles are the expectations of others with regard to a person's rights, obligations, and behaviors. The second, personality, is an individual process, although society still has a large effect, and combines the feelings, behaviors, intelligence, temperament, and interests of the individual.

This entry focuses on multiple examples of how the individual develops socially throughout the life course and examines all developmental stages from infancy to adulthood. Other issues that develop throughout the life course, including language development, moral development, and gender construction, are also discussed. Lastly, a discussion of problematic psychosocial development is included.

Infancy

Infancy marks the first contact with life outside the womb, and although it could be argued that psychosocial development begins prenatally, the research typically focuses on infancy as holding the beginning of its development; thus, this entry begins its discussion here. Infancy and childhood mark development patterns that have been found to have profound effects on later development and mental/physical well-being.

Perception of Faces

An infant's ability to perceive faces is the first sign that the child is developing a bond and interest with the human form. It was found through research that between 1 and 2.5 months, infants preferred a photo of

a schematic face over a photo of a scrambled face, which suggests that newborns have an innate preference for the human form. Recently it has been found that motion may have a large effect on this preference.

Social Smiling

Social smiling represents the first social activity of an individual. Smiling begins at 1 to 2.5 months of age but occurs at any external stimulation the child receives. It is not until 2.5 to 3 months that infants smile to reciprocate another's smile, or to have their smile reciprocated. This event typically begins to establish a new emotional bond between the child and the child's caregiver as well as show how the infant is beginning to create a sense of self and others in his or her environment, although this is a very basic sense.

Temperament

Temperament is an example of the nature side of psychosocial development, in that it is the innate characteristics or personality of a person. Temperament is thought to consist of activity level, approachability, adaptability, intensity of reaction, quality of mood, distractibility, attention span, and other qualities. Three types of infant temperaments have been named:

1. *Easy babies* are those who are adaptable, typically happy and playful, and not easily distracted.

2. *Difficult babies* are not adaptable, usually irritable, and easily distracted.

3. *Slow-to-warm-up babies* take a little more time to adapt, are less active but happy, and are usually not very distractible.

Temperament is thought to extend throughout the life course and remain fairly stable.

Attachment

Attachment not only is an important developmental milestone in infancy but also has deep effects on the rest of the life course. Harry Harlow, with the use of his monkey studies, was the first to realize that attachment was not created out of the need fulfillment of nourishment but was more strongly associated with feelings of comfort and safety. In his groundbreaking study Harlow placed two pseudo-mothers in to raise infant monkeys: one who provided nourishment but

was made of a bare wire body and one that did not provide nourishment but was wrapped in warm, comfortable blankets. The infants more often relied on and spent more time with the non-nourishing mothers.

Formation of Attachment

John Bowlby stated that attachment provide the infant with both safety and the ability to be autonomous, which allows him or her to have varied experiences. He proposed a four-phase model of attachment development.

1. *Preattachment* develops between birth and 6 weeks. The infant remains close to whomever provides food and comfort and does not feel distressed when left with strangers.

2. *Attachment in the making* occurs between 6 weeks and 6 to 8 months. The infant shows preference toward primary caregivers and slight distress toward strangers.

3. *Clear-cut attachment* takes place from 6–8 months to 18–24 months. The child expresses a great deal of stranger anxiety, and the primary caregiver becomes the secure base from which the child explores.

4. *Reciprocal relationship* is the final phase of attachment and occurs after 18 to 24 months. The child spends increased amounts of time away from his or her primary caregiver, and either the mother or the child interrupts the other to establish renewed contact every once in a while.

Types of Attachment

Four types of attachment exist and are defined based on the infant's response to a social experiment called the Strange Situation procedure. In this experiment the child is placed in a room to play. The child's primary caregiver is also in the room. A stranger enters the room and attempts to sit near the child while the primary caregiver exits the room. The child typically reacts by crying, and the stranger attempts to console the child. The caregiver then reenters the room and attempts to console the child if needed.

Securely attached children (65% of children) respond by playing contently when the caregiver is present, becoming upset when the caregiver leaves, and being comforted only when the caregiver returns, usually quieting very quickly. *Insecure anxious/avoidant* children (23% of children) respond by being indifferent when the caregiver is in the room, sometimes

becoming upset when the caregiver leaves, and if they are upset the stranger can console the child as often as the caregiver. *Insecure anxious/resistant* children appear anxious and stay close to the caregiver, are upset when the caregiver leaves, and are not consoled when the caregiver returns. These babies seem to desire the renewed contact while simultaneously pushing away from it. *Disorganized* children appear confused and fearful and seem disoriented; they react with little emotion.

It has been shown that caregivers of insecurely attached children are inconsistent in their response to their babies' cries and have little physical contact with them. Children with disorganized attachment are often neglected or abused. Studies have shown that attachment breeds healthy social competence, whereas unhealthy attachment can lead to aggression; unfulfilling, insecure adult relationships; and more serious concerns, including psychopathology.

Social Referencing

Social referencing refers to infants' use of their primary caregiver to keep them safe while they explore their social world. As infants encounter an unknown object or person, they will look to their reference point (primary caregiver) to notice their reaction to the same stimulus. If the caregiver looks wary, the infants will be hesitant or avoid the stimulus; if the caregiver looks content or neutral, the infants will be less likely to avoid the object. This becomes a primary method of communication between the caregiver and the infant.

Self-Concept

Infancy marks the beginning of the individual's understanding and recognition of himself or herself as a unique person. It is during this time that the individual first recognizes himself or herself in a mirror, something only humans and great apes accomplish, which occurs gradually over the first year of life.

End of Infancy

At the end of infancy, children begin to have less distress when separated from their caregiver, and their sense of themselves becomes distinctive. They begin to have standards similar to adults and recognize a violation of these standards, such as when they notice the difference between old, broken toys and new toys. It is at this transition that children begin to develop secondary emotions, like embarrassment, guilt, and envy, which begins to show their realization of how they have an effect on others in their social world.

Childhood

Childhood continues to increase the individual's level of self-understanding through engagement in friendship and peer groups as well as the changing relationship with his or her parents.

Self-Concept

As children enter school age, they begin to hold descriptions of themselves that include gender, friendship/group relations, as well as physical and psychological traits. This is the age at which they begin to compare themselves with other children, which shows their understandings of what is desirable in their social environment as they envy the traits of others. Their views of themselves are important because research has shown many beneficial outcomes for children who have positive self-concepts, including more confidence, autonomy, assertion, sociability, and a greater outlook on life. This age group begins to recognize and expect the stability of these traits and behaviors in themselves and others.

Self-Esteem

Self-esteem is how positively or negatively a person feels about himself or herself, and it is an important attribute of stable mental health. Positive self-esteem is enhanced by parental relationships that are warm and accepting while having clear and understood limits.

Self-Regulation

Along with their self-conceptions and self-esteem, children develop a sense of self-regulation, which means they learn how to act in socially appropriate ways and how to stop themselves before acting in inappropriate ways. This is a difficult task for the child to develop because they must learn how to stop behavior that has already begun. Play with other children has been shown to have a large effect on development of self-regulation because children learn how

to control their behavior and their emotions while watching and interacting with other children who may be more or less advanced than themselves.

Friendship and Peer Groups

Friendships and peer groups are an important aspect of psychosocial development because they provide the major social context in which youths interact with each other. As youths begin to separate from their familial relationships these social groups begin to establish greater influence and pressure on a child's life. These peer groups and friendships increase the development of important social skills. Playing games requires following rules and engaging in prescribed behaviors, both of which help children to understand larger and more complex rulings in their social environment. Children also begin to understand their place in the social world through their membership in cliques and crowds, which are the most common form of child and adolescent peer groups.

Peer Pressure

Although peer pressure is a common phenomenon within friendship and peer groups, it has been shown in research that youth are much more likely to go along with prosocial pressure than pressure to engage in antisocial behaviors. Peer pressure is believed to peak at the age of 15.

Influence of Parents

Two major changes occur as a child begins to associate more with peers and less with parents: (1) Parents' expectations and demands on domestic chores increase along with their demands and expectations on academic achievement, and (2) parental control is achieved less through direct methods and more through discussion and reasoning, sometimes with the inducement of guilt. The relationship and monitoring of parents can have a profound effect on peer relationships and later development.

Adolescence

Research has identified four major changes in the structure of adolescent social life:

1. A great deal more time is spent in peer relationships.

2. Guidance from adults lessens in quantity and directness.

3. There are large increases in mixed-sex interactions.

4. The participation in large social groups becomes important.

Beyond these changes in adolescents' social life, their search for identity and intimacy also have effects on their psychosocial development.

Identity

Identity is the individuality of a person and is composed of relatively stable personality traits that are especially prevalent in certain situations (role behaviors such as being a parent) and is thought of as the primary social concern for the adolescent. Although a development of identity occurs throughout the life course, most focus is placed on the adolescent years. According to Erik Erikson, a person must resolve his or her adolescent crisis before the person can confront the next stage crisis; thus, some sort of solution must be made with regard to identity, making it a rather important topic. Identity development is not a simple task; it is a process of understanding and synthesizing childhood experiences and memories, family history, present self-awareness and acceptance, and future goals or ideals. Beyond this internal process the person must also combine his or her relationships with persons and things external to the person.

Intimacy

Adolescents' search for intimacy includes their creation of new friendships and dating relationships, and these relationships increase adolescents' understanding and development of autonomy.

Gender Differences in Intimacy

As in childhood there are gender differences in relationship formation during adolescence. Females are more caring and emotionally open in relationships, whereas men are seen as being more independent, self-reliant, and closed off. Males are seen to be more assertive and domineering, while showing little emotion, and females are seen as more flirtatious and emotional. These gender differences peak in later adolescence, and by the age of 50, they are much more diminished.

Separation from Parents

During adolescence, while youths are attempting to forge their own identities, they begin to separate

from their parents. This causes a decrease in the amount of influence a parent has on the youth; however, research has shown that most adolescents maintain their parents' values, even though they may be more likely to verbally question them. Consistently throughout the research there is evidence that the most beneficial developmental outcomes stem from parents who take an authoritative stance toward their adolescent, meaning they exhibit warmth and respect, while holding clear, firm rules.

Adulthood

Adulthood is not always seen as increasing the development of a person; in fact, the early theorists stopped discussing development at the termination of adolescence. There are, however, a number of changes occurring throughout adulthood that affect our psychosocial development.

Early Adulthood

Early adulthood is considered a person's 20s and 30s and typically begins with a continued search for intimacy and progresses through the creation of a family unit.

Love and Loneliness

Erikson stated that this is a time in which the psychosocial crisis centers on intimacy versus isolation; another way to put this is love versus loneliness. One way in which young adults obtain this sense of love or intimacy is through their continued friendships, which tend to increase in the amount of emotional closeness, even in males but especially in females. Loneliness is said to be a large concern for college freshmen because it is usually associated with transitions in a person's life.

Types of Adult Status

Adults comprise six status types, which are not linear in dimension and thus do not form a hierarchical structure. *Single adults* are becoming an increasingly viable option for this age group, with many individuals remaining in this status throughout early adulthood. Many negative connotations of "singlehood" have been replaced with positive ones. *Cohabiting adults* is another status that has been on a recent rise. Research has shown that this type of relationship has either no effect or, if any, a negative

effect on marital satisfaction and divorce. *Married adults* have been on the decline since the 1960s but still represent many individuals in early adulthood. *Divorced adults* have started to slow their rapid increase in numbers in the recent years. This is a status that can have large emotional effects on future relationships and intimacy, although research suggests that, for most people, these effects are temporary. *Remarried adults* typically bring with them a combination of families, creating stepfamilies that can create a large adjustment for all members involved; it has also been found that only one third of all remarriages survive. Lastly, *gay and lesbian adults* have had much change in recent years concerning their right to be married. Research has shown that these adults are very similar to heterosexual adults. There are currently many misconceptions about gay and lesbian adults, including the belief that they are not good parents, when in fact they are similar in this regard, and the majority of children raised in these households grow up heterosexual.

Stages of Family Life Cycle

The family life cycle is seen as progressing through six linear stages: (1) leaving home and becoming a single adult, (2) joining of families through marriage, (3) becoming parents, (4) families with adolescents, (5) the midlife family, (6) the later life family.

Middle Adulthood

Middle adulthood creates a new set of challenges concerning psychosocial development when adults must deal with the transition of their children leaving home and in many cases their aging parents moving in.

Generativity Versus Stagnation

Erikson stated that the 40s and 50s were concerned with the crisis of generativity versus stagnation which meant that the person was left with the choice of helping the next generation, which is typically associated with leading a useful life, and not assisting the next generation and remaining stagnant.

Life Events Approach

Life events approach took a separate approach to Erikson's theory of psychosocial stages, saying that stage models are not as useful in later life because events sometimes happen out of order, or are skipped

in entirety, due to life events. This approach states that life events affect the individual not only due to the event itself, but also to how the person adapts to the event and the context of the event.

Marriage and Divorce

With regard to the life events approach, marriage and divorce can severely affect the development of a person in middle adulthood. It has been found that those who divorce in midlife are less likely than those in young adulthood to have a negative relationship with their divorced partner. It has also been found that divorce can have a positive effect on the female's emotional well-being; the opposite was found with men, who have more positive emotional states during marriage.

Parents and Children

Midlife is typically the period when children leave the home, creating the "empty nest." However, despite popular belief, this event typically increases the satisfaction of the parents, particularly dealing with marital satisfaction. This can also be a time in which the family must deal with bringing in the aging grandparents who may be in poor health, which creates what is called the *sandwich generation* (having both one's own children and one's parents to be responsible for).

Late Adulthood

Late adulthood is typically a time of life review, in which a person looks back onto his or her life in either satisfaction or disappointment.

Integrity Versus Despair

Erikson's final stage of psychosocial development concerned the crisis of integrity versus despair, which was reworked by Robert Peck to include three developmental tasks faced by older adults. (1) Differentiation versus role preoccupation concerns the redefinition of the individual's worth beyond his or her work roles now that the individual has retired. (2) Body transcendence versus body preoccupation concerns the fact that older adults must accept their declining physical well-being. (3) Finally, ego transcendence versus ego preoccupation concerns the need for older adults to accept the fact that they will die and that they have made an impact on the future of others.

Level of Activity and Social Support

An important component of older adulthood concerns remaining active. Studies have shown that older adults who remain active are more likely to look onto their lives with satisfaction. Social support from families or peer organizations has also been linked with longer lives and increased physical and mental health.

Gender Changes in Older Adulthood

A final change in the psychosocial development of an individual concerns gender. It has been found that gender differences all but disappear in older adulthood, and in some cases actually reverse. Males become more nurturing and sensitive while females are more assertive.

Life Course Psychosocial Developments

Some important changes in psychosocial development occur throughout the life span and thus are discussed separately.

Language Development

One of the most important developments in a psychosocial world is the ability to vocalize our thoughts and emotions. Language development begins in infancy as the child begins to hear speech in his or her social world, and continues through childhood, all the way into adulthood as it becomes more complex and varied. At approximately 9 months, children utter their first words, but their ability to use words to command attention does not exist until around 12 months. At 18 months they are forming two-word sentences with a large increase in vocabulary, and finally at 24 to 30 months, they are beginning to take the listener's perspective into account and responds to indirect requests. This development increases the individual's place in the social world dramatically because the individual can now be a direct actor in his or her social world and achieve a numerous variety of goals.

Moral Development

The development of morals is a rather complex topic that begins in childhood and lasts into adulthood with few adults obtaining the highest levels of morality. Lawrence Kohlberg stated that the early stages of morality are considered preconventional reasoning

that typically ties together morality and punishment, which stems from an external authority. A child in this stage believes in imminent justice in that if a rule is broken, a punishment will occur immediately regardless of whether there is a witness. The later stages of postconventional reasoning consist of social contracts and individual rights, as well as universal ethical principles of universal human rights.

Gender Construction

Gender is defined as the social construction of being male or female. An individual typically has an established gender by the age of 3. Gender roles are socialized norms of how a person should act and feel dependent on his or her gender. Most children develop gender roles initially based on those of their parents; however, these roles develop and are rewarded or punished by peers throughout life.

Defining an Identification

Defining one's identification may be confused with defining an identity, although identification does create strong roots for a person's identity later in life. Identification is the way a person attempts to look, behave, and feel, as well as who the person identifies with in his or her social world. Individuals also identify through the creation of gender roles which was discussed earlier.

Problematic Psychosocial Development

Not all individuals develop in the same way concerning psychosocial development. Some develop in a problematic way, typically because of some deficit or excess in their life course.

Possible Causes of Problematic Development

Some links that have been found associated with problematic development include negative or absent attachment, child abuse, and negative or inconsistent parenting styles. Other important causes outside the parent–child relationship are too much television, which can expose the child to increased violence and diminish creativity and verbal skills, and large work schedules of adolescents, although up to 20 hours per week of part-time work during high school has been shown not to have a negative effect on development.

Psychological Disorders Associated With Problematic Development
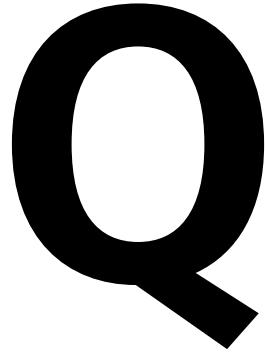
In some cases the psychosocial development process can have large effects on an individual's mental health and create psychological disorders, including substance use, antisocial behavior, depression, suicide, autism spectrum disorders, and many others.

Matthew J. Davis

See also Attachment Disorder; Autism Spectrum Disorders; Cliques; Erikson's Theory of Psychosocial Development

Further Readings

Cole, M., Cole, S. R., & Lightfoot, C. (2005). *The development of children* (5th ed.). New York: Worth.

Santrock, J. W. (2002). *Life-span development.* New York: McGraw-Hill.

Steinberg, L. (2005). *Adolescence* (7th ed.). New York: McGraw-Hill.

Q

Torture numbers and they'll confess to anything.

—Gregg Easterbrook

QUALITATIVE RESEARCH METHODS

Qualitative methods are used typically in research projects that take a human-focused perspective in the design and implementation of the investigation. *Qualitative research*, as a paradigm or worldview, is designed to explore the human elements of a topic under investigation; in this context, *qualitative methods* (such as in-depth interviewing or ethnographic observation) are used to examine how individuals see and experience the world around them. Researchers typically talk to people directly or observe their behaviors in various contexts to understand what those individuals view as important about a particular phenomenon. Although qualitative research is often described in opposition to quantitative research, and although the goals and intentions of these paradigms are quite different, many scholars are now using mixed and multimethod approaches and engaging in interdisciplinary research. In these projects, both qualitative and quantitative paradigms are used to more fully investigate a research problem. Indeed, many researchers now use qualitative methods to inform future quantitative research designs (and vice versa) or to add a richness of experience to an investigation by combining qualitative interviews or diaries with findings from large-scale questionnaires or experimental interventions.

The Qualitative Research Paradigm

The use of qualitative methods has a rich history in the social sciences, health sciences, and humanities. In the field of education, many research problems are best addressed by qualitative approaches, so the use of qualitative methods is very much the norm. In educational psychology, case studies, interviews, and other techniques have been used recently to explore a range of research topics. Although Gary Shank noted in 1994 that qualitative research had "made very little headway" (p. 340) in educational psychology at the time, the rise in the number of qualitative projects over the past decade points to a positive shift toward including more qualitative approaches. The rise of qualitative methods in the field of psychology itself, as noted in Sean Kidd's review of core journals, also points to the value of using qualitative approaches to investigate problems in disciplines that have been focused traditionally on quantitatively oriented research practices.

Research of high quality, regardless of paradigm, demands that appropriate methods are used to address the research problem at hand. To select appropriate methods, researchers must first understand the types of data they will obtain with each approach. To assess student satisfaction with in-school counseling, for example, a researcher should start by asking, "What do I want to know?" and "What method will allow me to

address this question?" Unfortunately, new scholars (or those unfamiliar with the intent of qualitative methods) often make the mistake of starting a project by saying, "I want to use an online questionnaire" or "I want to design an experiment," even when that method may not be the best one to provide data for a particular research problem. Understanding the intended goals of qualitative research is essential to selecting appropriate methods and to assessing the results.

Qualitative methods are best for addressing many of the "why" questions that education researchers have in mind when they develop their projects. Whereas quantitative approaches are appropriate for examining "who" has engaged in a behavior or "what" has happened (e.g., How many students approached school counselors for help last year? For what types of problems did they seek help?), and whereas experiments can test particular interventions, these techniques are not designed to explain *why* certain behaviors occur. Quantitative approaches are best used to document characteristics of the world around us, that is, what we see and what the implications are when we test hypotheses related to specific phenomena. As Ted Palys and Chris Atchison note in their text, *Research Decisions*, quantitative researchers prefer the deductive method, whereby a researcher deduces a hypothesis from a theory, gathers data to test the hypothesis, and then revises (or discards) the theory or looks for another situation in which to test the theory again; here, the "true experiment" is the method of choice, as this allows the effects of certain variables to be assessed while all other influences are held constant. Qualitative approaches, on the other hand, are typically used to explore new phenomena and to capture individuals' thoughts, feelings, or interpretations of meaning and process. Many qualitative texts (e.g., David Silverman's book, *Doing Qualitative Research: A Practical Handbook*), provide an overview of the preferences of qualitative researchers: to gather qualitative data that are naturally occurring, to explore meanings rather than behaviors, to reject the natural science model, and to craft studies that are inductive and hypothesis generating rather than ones that involve hypothesis testing.

The Nature of Qualitative Data

As qualitative inquiry allows a researcher to examine a topic in great depth, the data gathered tend to be very rich in scope. A researcher may compile dozens of pages of field notes in completing an ethnographic observation, or capture many hours of video during focus group interactions. Data collection and analysis typically occur over many weeks or months, even in one setting or with only one small group of participants as the focus of the investigation. The depth of the investigation makes qualitative research both costly and time consuming. The sample sizes used in qualitative research are usually quite small (in some cases, only one participant or setting may be used; in others, 20 to 30 participants may be appropriate) and may involve relatively homogenous groups. However, as generalizability is not the intent of such projects, small samples that are carefully and deliberately chosen (e.g., using maximum variation sampling techniques) provide the best data possible for this type of work. Some researchers will also use multiple qualitative methods to explore one research problem (in a process known as *triangulation*), which can extend the time and cost involved in completing a single project. Dissemination of qualitative results can extend across many years, as the analysis of the rich data resulting from these types of projects can lead to multiple publications covering a broad range of research questions. Also, many researchers engage in projects that build on the results of prior studies, which can extend the length and breadth of a single project over many years and across various contexts or populations.

In delineating the nature of qualitative research, it is also important to note the distinction between the gathering of qualitative "data" and the use of a qualitative perspective or "paradigm" in research design and implementation. Open-ended questions used as part of a questionnaire, for example, allow respondents to provide their opinions and attitudes about a particular topic or event. Gathering such data is often done alongside quantitative measures (e.g., Likert scales) to elicit information that cannot be captured by closed questions. However, including open-ended questions does not make a research method inherently "qualitative." Rather, qualitative research reflects a particular orientation to research that informs all phases of the research design—from the study's conception, through data gathering, analysis, and writing. Many qualitative research projects are designed by scholars who see knowledge as socially constructed and inextricably linked to individuals' backgrounds, personal histories, cultural place, and other contextual elements that define the human experience. Whereas quantitative researchers strive to eliminate these contextual

elements so that they will not "contaminate" the variables under study (to maintain objectivity and reduce bias), qualitative researchers embrace these elements, to examine phenomena in light of these contextual characteristics. Qualitative projects purposefully embrace bias and subjectivity, which, in turn, often become a focus of the investigation.

Qualitative Methods at a Glance

Qualitative researchers, across disciplines, use a vast array of methods in their work. Indeed, in education alone, researchers use interviews, focus groups, participant observation, diaries, think alouds, and a host of other qualitative techniques to examine students', teachers', and parents' experiences of the school environment. New technologies have also recently changed the research landscape; many qualitative researchers now use such tools as videoconferencing or digital video to capture data or use the Internet as a site of qualitative inquiry. Although the following sections provide brief descriptions of a few methods commonly used by education researchers, this is by no means an exhaustive list.

Interviews

In-depth, qualitative interviewing is one of the most commonly used methods. These interviews allow researchers to examine issues, at length, from the interviewee's personal perspective. The data gathered typically consist of verbatim responses to the interviewer's questions, which are designed to elicit opinions, feelings, attitudes, descriptions of personal behaviors, and other elements related to the research problem. The method may use a semistructured approach (e.g., open-ended questions are used to capture participants' thoughts of a variety of topics), or they may be quite unstructured (e.g., participants are encouraged to explore a topic, such as their personal life story, in their own way). Qualitative researchers (such as Irving Seidman) note that the purpose of this type of research is to understand individuals' experiences and the meanings they make of those experiences and to put their behaviors in context in order to understand the actions they undertake. Semistructured interviews typically last from 60 to 90 minutes, although the length and the number of interview sessions will vary depending on the scope of the project, the availability of participants, other methods used in

the project, and so forth. An unstructured interview may last over a number of hours or across many days or weeks, particularly in ethnographic studies. Common themes and patterns are identified in the data gathered, with saturation of themes (resulting in transferable data) occurring with 15 to 18 participants. Increasing the number of interviewees is one way to enhance rigor in data collection. However, anomalies in the data—where individuals discuss an experience unlike that of other interviewees—are also extremely valuable, as they can highlight very specific needs and points of conflict.

Focus Groups

Focus groups also fall into the qualitative interviewing category, but in this case, the interviews occur with groups of individuals (typically between five to eight people, with one or more groups in total). The discussions are generally focused on a particular issue, though many groups are designed to engage in open, exploratory dialogue. This method can also be used to assess the value of a product, service, or tool (e.g., in Web site usability studies) or to gauge individuals' attitudes, opinions, or beliefs about a social topic (e.g., the influence of the new media on political decisions). Group interviews can be more challenging to conduct than individual interviews because group dynamics need to be managed (e.g., to ensure that all group members have a chance to speak) and because special equipment (e.g., multiple video cameras) is often used. Focus group interviews are best run by a trained facilitator; they often require a more formal setting (such as a boardroom) and may take more time to coordinate and run.

Observation

Observing human behavior in natural settings (e.g., watching students as they interact in a classroom) allows a researcher to capture valuable data that cannot be gathered in any other way. Observational techniques may be covert or overt; researchers may be either distanced from, or directly engaged in (i.e., participant observation), the daily activities under investigation. Researchers employing observational methods not only document details about the individuals in the setting (e.g., a school), but they also examine the physical (e.g., desk placement, computer lab setup) and organizational (e.g., school administration) structures

within the setting. Data collection may involve one or many sites and typically extends over a long period of time.

Personal Journals and Diaries

Asking individuals to document their activities (e.g., how, and for what purposes, students use the Internet to seek counseling support) is an effective way to examine behavior without relying on memory (as with interviews, where people are asked to recall past behaviors). Diaries allow individuals to document quantitative elements of their activities (e.g., how often they use the Internet) and qualitative opinions or attitudes about their experiences. Participants typically need instruction about how much to write, how often, and on what topics, but these journals can produce much more detailed accounts than is possible with other methods. Individuals may keep diaries for a week or more and may write on a variety of topics, which can then be further explored through other, complementary methods (e.g., personal interviews). Many researchers also now provide digital audio or video recorders to participants, particularly where writing a diary may prove difficult, inappropriate, or time-consuming.

Critiques of Qualitative Methods

Although many texts and journal articles on research methods articulate the varied goals and intentions of the various research paradigms, the use of qualitative methods continues to be controversial in some disciplines. Some scholars (particularly in biomedical sciences and in very quantitatively oriented disciplines) dismiss qualitative research outright, arguing that these studies lack rigor in design and execution, are biased, or present only anecdotal findings. Silverman notes, for example, that questions are often posed about a qualitative study's reliability or validity when these quantitative measures of rigor are simply inappropriate for judging qualitative research. Other methodologists note that it falls to the researchers themselves to defend their work in light of charges from those who sit outside of the qualitative paradigm. Yvonna Lincoln and Egon Guba, in their classic work *Naturalistic Inquiry*, noted that qualitative researchers learn quickly that they will face charges of engaging in "sloppy" research or "merely subjective" observations (p. 289). Unfortunately, the

concerns raised in Lincoln and Guba's text about unwarranted criticisms being lodged against qualitative researchers remain a concern in some disciplines, today, more than 20 years after the publication of their book.

However, as qualitative methods have been employed for decades across different disciplines (e.g., sociology, anthropology, social work, information science, education, nursing), and as researchers have published extensively on qualitative methods development (e.g., in the *International Journal of Qualitative Methods*), one might reasonably argue that these types of critiques are grounded more in ignorance or misunderstanding than in substance. Rather, the criteria used to assess the rigor of qualitative studies are quite different from those used in quantitative studies. In quantitative research, validity, reliability, generalizability, and objectivity are the markers of quality work. Quantitative results are intended to be free from bias, to be replicable across contexts, and to generalize from the sample under study to the full target population (e.g., to all undergraduate students requiring counseling services). Qualitative research is assessed using the markers of credibility, transferability, dependability, and confirmability, including techniques such as prolonged engagement, persistent observation, triangulation of methods, negative case analysis, peer debriefing, member checks, or a combination of these or other techniques to ensure a study is of high quality. Lincoln and Guba provide an excellent overview of the main tenets of these approaches to rigor in qualitative research.

Qualitative Methods and Research Ethics

It is also important to note that qualitative methods bring with them very specific issues and challenges that must be addressed in researchers' applications for ethics approval. Many scholars (e.g., Will van den Hoonaard) have published texts that examine the research ethics associated with a variety of qualitative methods and point to some of the common issues that researchers must address. In ethnographic studies, for example, where a researcher conducts a project involving participant observation, confidentiality and informed consent procedures may need particular attention (especially if the researcher's role is not disclosed to members of the community under study). With Internet studies involving private (subscription

only) chat sites, the process for obtaining and maintaining consent may be problematic, due to the ways that participants sign in and out at will. Using video cameras in public spaces may also pose a challenge to qualitative researchers, as individuals' faces or voices may be identifiable; in these types of projects, a researcher may seek approval to disclose individuals' identities (with their informed consent) or may need to alter the digital record to preserve participants' anonymity. All of these ethics-related issues and choices are governed by ethics guidelines (which vary by jurisdiction), as well as by disciplinary norms and strategies developed over many decades. It is important that researchers who plan to use particular methods, techniques, and technologies refer to published texts or seek advice from established qualitative scholars in the implementation of appropriate ethics practices to suit those approaches. Recently, ethics boards have been criticized for requiring researchers to impose inappropriate guidelines or ethics-related practices in their qualitative studies (see, e.g., the *Giving Voice to the Spectrum* report published in 2004 by Canada's Social Sciences and Humanities Research Ethics Special Working Committee). It is important that researchers engaged in qualitative work educate themselves on the ethics implications of this type of work and demand that ethics review panels implement appropriate procedures and offer appropriate advice on the design of these types of projects.

Conclusion

Qualitative methods have an important, valuable place in research in educational psychology. However, to select appropriate methods, regardless of paradigm, researchers must focus on the research problem at hand and choose methods that will provide evidence of high quality to address that problem. When projects are designed based on a preference for particular methods (e.g., when a researcher wants to create an online questionnaire because of his or her interest in new technology) or based on a belief that a particular paradigm is better than another (e.g., that quantitative data are inherently more valuable than qualitative data), the overall worth of a study comes into question. As long as particular methods (and paradigms) are privileged over others, with no regard for context or for the desired outcome for the research problem itself, the quantitative versus qualitative divide that continues to drive some

educational research approaches will only gain more ground. Qualitative and quantitative approaches are complementary paradigms that examine the world from two very different vantage points, and each has its own important place in advancing knowledge in the field of education. Indeed, projects that are designed to suit the intended goals of a paradigm or that embrace both sets of methods (if appropriate) go a long way in illuminating understanding of the diverse range of human behaviors.

Lisa M. Given

See also Case Studies; Ethics and Research; Ethnography; Longitudinal Research; Naturalistic Observation

Further Readings

Kidd, S. A. (2002). The role of qualitative research in psychological journals. *Psychological Methods*, 7, 126–138.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage.

Mertens, D. M. (2005). *Research and evaluation in education and psychology: Integrating diversity with quantitative, qualitative and mixed methods* (2nd ed.). Thousand Oaks, CA: Sage.

Palys, T., & Atchison, C. (2008). *Research decisions: Quantitative and qualitative perspectives* (4th ed.). Toronto, ON: Thomson Nelson.

Patton, M. Q. (2001). *Qualitative evaluation and research methods* (3rd ed.). Thousand Oaks, CA: Sage.

Seidman, I. (2006). *Interviewing as qualitative research: A guide for researchers in education and the social sciences* (3rd ed.). New York: Teachers College Press.

Shank, G. (1994). Shaping qualitative research in educational psychology. *Contemporary Educational Psychology*, 19, 340–359.

Shank, G. D. (2006). *Qualitative research: A personal skills approach* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.

Silverman, D. (2000). *Doing qualitative research: A practical handbook*. London: Sage.

Social Sciences and Humanities Research Ethics Special Working Committee. (2004). *Giving voice to the spectrum: Report of the Social Sciences and Humanities Research Ethics Special Working Committee*. Ottawa, ON: Interagency Advisory Panel and Secretariat on Research Ethics. Retrieved June 8, 2007, from http://www.pre.ethics.gc.ca/english/workgroups/sshwc/reporttopre.cfm

Social Sciences and Humanities Research Ethics Special Working Committee. (2007). *Qualitative research in the context of the Tri-Council policy statement—Ethical conduct for research involving humans: A discussion paper*. Ottawa, ON: Interagency Advisory Panel and

Secretariat on Research Ethics. Retrieved June 8, 2007, from http://www.pre.ethics.gc.ca/english/workgroups/sshwc/consultation07.cfm

van den Hoonaard, W. (Ed.). (2002). *Walking the tightrope: Ethical issues for qualitative researchers.* Toronto, ON: University of Toronto Press.

Quantitative Research Methods

Educational professionals are concerned with learning, achievement motivation, human growth and development, and human behavior. Consequently, measuring change in these areas is of great interest and need to educational researchers. The purpose of quantitative research is to find solutions to problems worthy of investigation. Results of quantitative research provide objective evidence as a basis for decision making. These decisions may affect individuals, organized programs of study, or entire school systems. For example, new teaching methods may be implemented, new programs established, curricula modified, new programs implemented, or new evaluation procedures used based on the results of quantitative research. Over time, results of quantitative research can influence an entire field of study; such was the case with Piaget's research in child development.

Quantitative research methods allow researchers to measure certain behaviors or phenomena. Data are represented in the form of numbers. Quantitative research methods require the scientific method of investigation. Measurement is necessary if the scientific method is to be used. The scientific method involves an empirical or theoretical basis for the investigation of populations and samples. Hypotheses must be formulated, observable and measurable data must be gathered, and appropriate mathematical procedures must be used for the statistical analyses required for hypothesis testing. If researchers are to ascertain relationships between variables and differences among variables, data must be provided to support objective, valid, and reliable conclusions. The conclusions should be based on predetermined hypotheses and statistical testing. Data may be collected in various ways, such as via the Internet, surveys, tests, inventories, or written documents.

There are many quantitative methods that may be applied to research in the behavioral and social sciences. Both descriptive and inferential statistics are used in quantitative research. Quantitative methods depend on the design of the study (experimental, quasi-experimental, nonexperimental). Study design takes into account all those elements that surround the plan for the investigation, for example, research question or problem statement, research objectives, operational definitions, scope of inferences to be made, assumptions and limitations of the study, independent and dependent variables, treatment and controls, instrumentation, systematic data collection actions, statistical analysis, time lines, and reporting procedures.

Stages

Problem Statement

First, an empirical or theoretical basis for the research problem should be established. This basis may emanate from personal experiences or established theory relevant to the study. From this basis, the researcher may formulate a research question or problem statement. Even though a study may have several objectives, a clear research problem statement or research question should be stated. Often these are teacher observations of classroom behaviors. When teachers formulate theories and engage in research, it is called *action research*. The research problem, whether anchored in theory or an outgrowth of personal observation, provides direction for the methods to be used.

Operational Definitions

Operational definitions describe the meaning of specific terms used in a study. They specify the procedures or operations to be followed in producing or measuring complex constructs that hold different meanings for different people (e.g., beauty, intelligence, work values). For example, intelligence may be defined for research purposes by scores on the Stanford–Binet Intelligence Scale.

Population and Sample

Quantitative methods include the target group (population) to which the researcher wishes to generalize and the group from which data are collected (sample). Early in the planning phase, the researcher should determine the scope of inference for results of the study. The scope of inference pertains to populations of interest, procedures used to select the sample(s), method for assigning subjects to groups, and

the type of statistical analysis to be employed. All of these elements in a study affect the extent to which conclusions can be inferred beyond the study unit(s). In experimental research, a study unit is that which receives the treatment, such as people, animals, or plants. When the whole population is used in a study, no inferences need to be made. Usually one or more samples from the population are used in a study.

Hypothesis Testing

Complex questions to compare responses of two or more groups or show relationships between two or more variables are best answered by hypothesis testing. A hypothesis is a statement about the researcher's expectations about a relationship between variables to be studied. These statements may be written in the alternative or null form. A directional alternative hypothesis states the researcher's predicted direction of change, difference between two or more sample means, or relationship among variables. An example of a directional alternative hypothesis is as follows:

> Third-grade students who use reading comprehension strategies will score higher on the State Achievement Test than their counterparts who do not use reading comprehension strategies.

A nondirectional alternative hypothesis states the researcher's predictions without giving the direction of the difference. For example:

> There will be a difference in the scores on the State Achievement Test for third-grade students who use reading comprehension strategies and those who do not.

In quantitative research, the null hypothesis is the hypothesis of interest. Stated in the null form, hypotheses can be tested for statistically significant differences between groups on the dependent variable(s) or statistically significant relationships between and among variables. The null hypothesis uses the form of "no difference" or "no relationship." An example of a null hypothesis is as follows:

> There will be no difference in the scores on the State Achievement Test for third-grade students who use reading comprehension strategies and those who do not.

Hypothesis testing requires inferential statistical procedures. It is important that hypotheses to be tested be stated in the null form, because the interpretation of the results of inferential statistics is based on probability. Testing the null hypothesis allows researchers to test whether differences in observed scores are real or due to chance or error; thus, the null hypothesis can be rejected or retained.

Data Organization and Preparation

The organization and preparation stage is important and involves several preliminary steps. Survey forms, inventories, tests, and other data collection instruments returned by participants should be checked for completeness to assure that there are no missing data points. If the data set is small, the researcher can visually check the returned instruments. For larger data sets, a statistical program or spreadsheet is recommended. Missing data may be due to participants' oversight, unwillingness to answer certain questions, or lack of knowledge about the question asked. The researcher should consider carefully the procedures to be used to handle missing data, and the procedures used should be stated in the final report.

John Tukey suggested that exploratory data analysis be conducted using graphical techniques such as plots and data summaries in order to take a preliminary look at the data before the actual analysis is conducted with more formal methods. Exploratory analysis provides insight into the data set and the underlying structure of the data. The existence of outliers, data entry errors, unexpected or interesting patterns in the data set, and assumptions of the planned analysis, such as homogeneity of variance and normality of error for an analysis-of-variance statistical procedure, can be checked with exploratory procedures. Inferential statistics for tests of hypotheses should be conducted only after all preliminary analyses are completed.

Data Analyses

It is important that the researcher use the appropriate statistical tests for data analyses; otherwise, the results could lead to incorrect conclusions. Sometimes it may not be clear which test is the best for a particular study. Some important considerations are

1. type of research questions to be answered or hypotheses to be tested,

2. number of independent and dependent variables in the study,

3. number of covariates in the study,

4. scale of the measurement instrument(s) used to collect the data (nominal, ordinal, interval, ratio), and

5. type of distribution (normal or non-normal).

Examples of statistical procedures commonly used in educational research are *t*-test for independent samples, analysis of variance, analysis of covariance, multivariate analysis of variance, Pearson product moment correlation, Mann–Whitney U test, Kruskal–Wallis test, and Friedman's chi-square test. Many statistics textbooks, reference materials, and Web sites list decision tree diagrams or tables that provide guidance for selecting the appropriate statistical procedure.

Results Interpretation and Conclusions

The level of statistical significance that the researcher sets for a study is closely related to hypothesis testing. This is called the alpha level. It is the level of probability that indicates the maximum risk a researcher is willing to take that observed differences are due to chance. The alpha level may be set at .01, meaning that 1 out of 100 times the results will be due to chance; more commonly, the alpha level is set at .05, meaning that 5 out of 100 times the results will be due to chance. Using the .01 and .05 alpha levels means that if the null hypothesis is true, a low probability value will be observed. Alpha levels are often depicted on the normal curve as the critical region, and the researcher must reject the null hypothesis if the data fall into the predetermined critical region. When this occurs, the researcher must conclude that the findings are statistically significant. If the researcher rejects a true null hypothesis (there is, in fact, no difference between the means), a Type I error has occurred. Essentially, the researcher is saying there is a difference when there is none. Type I errors can have serious consequences in educational settings; for example, if a program or entire school were to revise a curriculum based on the results of such a study, vast resources will have been wasted. On the other hand, if a researcher fails to reject a false null (there is in fact a difference), a Type II error has occurred. In this case, the researcher is saying there is no difference when in reality a difference exists. The power in hypothesis testing is the probability of correctly rejecting a false null hypothesis. The cost of committing a Type I or Type II error rests with the consequences of the decisions made as a result of the test.

Tests of statistical significance provide information on whether to reject or fail to reject the null hypothesis; however, an effect size (R^2, eta^2, Phi, or Cohen's *D*) should be calculated to identify the strength of the conclusions about differences in means or relationships among variables.

Experimental Designs

True experimental designs are the most rigorous quantitative research methods in that they allow the researcher to have full control over the experiment and to assign subjects randomly to groups. Full control and randomization strengthen internal validity; however, external validity may be compromised. Experimental designs are useful for establishing cause-and-effect relationships among variables. One or more variables are systematically manipulated so that effects of the manipulated variables on two or more groups or individuals can be observed. The variable being manipulated is called an *independent variable* (treatment) and the observed variable (measured outcome) is called a *dependent variable*. True experimental designs are difficult to conduct in educational settings, because it is unrealistic to expect a classroom to mimic a laboratory. True experimental designs satisfy the following criteria: randomization, experimental control(s), experimental treatment(s), experimental and control group(s), standardized test instruments, and basic patterns of research designs.

Randomization

Random assignment of subjects to groups helps to control for bias in the selection process and assures that groups are similar on all important variables that may affect the outcome of the study before the treatment is administered. Probability sampling procedures assure that each participant in a study has an independent and equal chance of being selected for any group.

Experimental Control

The researcher should attempt to hold all variables constant that might influence the outcome of the study by allowing only the dependent variable to vary based on participants' responses to the treatment.

Researchers can control for time of day for a particular class, method of instruction, and same teacher in both classes. Achievement motivation, ability level, and satisfaction with school would be difficult to control, yet these variables could influence the outcomes of a study. It is difficult to impossible to have perfect control in educational studies because there are a host of extraneous variables that could affect a participant's response on the dependent variable. Pretesting, matching, blocking, and using covariates are common controls.

Experimental Treatment

One or more interventions or treatments may be manipulated in an experiment. New teaching methods, alternative testing methods, and different curricula are examples of treatments used in educational research.

Experimental and Control Groups

A researcher could compare reading achievement of two groups of students. The manipulated variable would be method of instruction (whole language or traditional), and the dependent variable would be reading achievement scores. The experimental group is taught reading by the whole language approach, and the control group is taught reading by the traditional approach. After the experiment, both groups would be tested on the same standardized test. Any change in reading achievement scores may be tentatively attributed to method of instruction.

Standardized Test Instruments

A test that meets certain standards or criteria for technical adequacy in construction, administration, and use is said to be standardized. A standardized test provides clear directions for administration and scoring, so that repeated administrations and scoring of the test are carried out systematically. The American Psychological Association, American Educational Research Association, and National Council on Measurement in Education have established standards that address professional issues and technical characteristics of standardized test development and use in the social sciences. Test validity, reliability, and established norms are major requirements of the standards.

Basic Patterns

Donald Campbell and Julian Stanley suggested three basic patterns of true experimental research designs: pretest-posttest control group design, posttest-only control group design, and Solomon four-group design.

Pretest-Posttest Control Group Design

A pretest is administered to a control group and an experimental group prior to the administration of the treatment. After the experiment, a posttest is administered to both groups, and gain scores from the pretest to the posttest may be compared. Statistically significant differences between gain score means may be computed using a *t*-test for independent samples, if only two groups are involved.

Posttest-Only Control Group Design

When pretesting is not possible or desirable, the posttest-only control group design may be used. Random assignment of subjects to groups serves to assure equality of groups. Gain scores cannot be computed; otherwise, statistical analysis for the posttest-only design is the same as that for the pretest-posttest control group design.

Solomon Four-Group Design

As the name implies, the design requires four groups. The design is configured so that one experimental group receives the pretest and treatment: One control group receives the pretest but no treatment; one control group receives the treatment but no pretest; and one control group receives neither pretest nor treatment. All groups receive the posttest. The Solomon four-group design is a combination of the pretest-posttest control group design and the posttest-only control group design. Essentially, the design requires the conduct of two experiments: one with pretests and one without pretests. The design provides more rigorous control over extraneous variables and greater generalizability of results than either of the previous designs. Because the design does not provide four complete measures for each of the groups, statistical analysis is performed on the posttest scores using a two-way (2×2) analysis of variance procedure. Thus, the researcher will be able to ascertain the main effects of the treatment, main effects of pretesting, and the interaction of pretesting with the treatment.

Quasi-Experimental Designs

Random assignment of subjects to groups may not be possible because it may disrupt school schedules or it may limit the number of classes in the discipline to be studied. In such cases, quasi-experimental designs are useful. Much of the research in education and psychology is conducted in the field or in classroom settings using intact groups. In these cases, researchers randomly assign treatments to the non-randomly selected subjects. The lack of full control and nonrandom assignment of subjects to groups pose threats to the internal and external validity of quasi-experimental designs. Matching may be used to control for the lack of randomization. In matching, researchers try to select groups that are as similar as possible on all important variables that may affect the outcomes of a study. Pretests are also recommended to control for lack of randomization. Similar scores on a pretest administered to all groups indicate that the groups were matched adequately. There is no doubt that quasi-experiments are weaker than true experiments for making causal inferences; however, information resulting from quasi-experiments is usually better than no information at all.

Nonrandomized Control Group, Pretest-Posttest Design

In an ideal experiment, subjects are assigned randomly to groups; however, random assignment is not always possible in school settings because of class scheduling conflicts. Rearranging students' schedules disrupts their classes, and obtaining permission from administrators to collect data may be difficult. Administrators are more likely to agree to a researcher entering two or more intact classes to collect data than to rearranging the class schedule of several groups of students. In other words, using intact groups for data collection is not only more convenient for the researcher, but it is also more palatable to the administration. For these reasons, the nonrandomized control group, pretest-posttest design is a popular choice for education researchers. A random procedure is used to determine which group or groups will be the control and which will be the experimental. In these situations, the researcher cannot assign subjects randomly to groups, so preexperimental sampling equivalence cannot be ensured. Similar scores on a pretest for the experimental and control groups denote a greater degree of preexperimental equivalency. After the pretest, the treatment is introduced, and both groups are administered the same posttest. Differences in pre- and posttest mean scores for the control and experimental groups are more credible when pretest means are similar. Statistical tests that compare mean scores are appropriate for the nonrandomized control group, pretest-posttest design when pretest scores indicate group equivalency. When results of the pretest reveal that the scores are not similar, more involved statistical procedures are required to address the measurement error on the pretest.

Time Series

Periodic measures are taken on one group at different intervals over an extended time period. Essentially, a time series design involves a series of pre- and posttest measures. A treatment is introduced, and its effects are assessed based on the stability of the repeated measures and differences from one measure to another.

Single-Subject Designs

The researcher investigates a single behavior, or a limited number of behaviors, of one participant by making initial observations, administering a treatment, and observing and measuring behavior after the treatment to study the effects. Pretreatment observations are necessary to establish baseline data. A simple single-subject design is the AB design where pretreatment observations and measures (A) are made to establish baseline data, and then treatment (B) is administered, and observations and measures (A) are repeated to ascertain changes in the behavior. Monitoring of the treatment usually lasts until it is apparent that the treatment has had some effect on behavior. The sequence of observations and measures (A) and treatment (B), followed by prolonged observations and measures (A), may be repeated as in the ABA and ABAB designs.

Factorial Designs

Many research questions in education require the investigation of more than one independent variable. Factorial designs permit the investigation of the effect of one independent variable on the dependent variable while ignoring other independent variables (main effect) and the combined influence (interaction) between two

or more independent variables. Full control and randomization are difficult at best in complex factorial designs. In fact, some factorial designs may require nonrandom assignment of subjects to groups, such as when subjects are placed in groups based on some predetermined criterion, such as ability level. Notation for factorial designs is usually written to identify the number of levels involved in each independent variable. For example, a "two by three" (2×3) factorial design would have two independent variables (factors) with two levels on the first independent variable and three levels on the second independent variable. All groups receive each level on each of the independent variables. A 2×3 design would require six groups. An example of a 2×3 factorial design is a study of method of teaching reading (whole language vs. traditional) and motivation (high, medium, low). Three groups would receive whole language reading instruction, and three groups would receive the traditional reading instruction. The dependent variable would be posttest scores on a standardized reading test. Interpretations of the analysis are more complex for factorial designs than for designs discussed previously. Analysis-of-variance statistical procedures are used to analyze data for factorial designs.

Nonexperimental Designs

It is not always possible, practical, or ethical to manipulate the research variables of interest. A parent's involvement in his or her child's education, an individual's level of job satisfaction, or whether or not a student has transferred from one school to another are variables that cannot be manipulated by the researcher. In these cases, the variable has already occurred. For such studies, nonexperimental methods are used. Three common nonexperimental methods of quantitative research are causal comparative (ex post facto), survey, and correlation.

Causal Comparative

Causal comparative studies do not permit the researcher to control for extraneous variables or for the randomization or manipulation of independent variables as in experimental studies. Changes in the independent variable have already occurred before the study is conducted; thus, these studies are also known as *ex post facto*, meaning "from after the fact." Research of interest may be on differences in subjects

on dependent variables when they have known differences on independent variables or the extent to which subjects differ on independent variables when they have known differences on the dependent variables. These studies require that the researcher select participants who already have or do not have the attribute variables being measured.

Survey

Survey research is a kind of descriptive study that is widely used in applied social sciences research. Sample surveys, unlike census surveys that query the entire population (as in the U.S. Census), collect data from a sample of individuals thought to be representative of the larger (target) population to which the researcher wants to generalize. Surveys are used to collect data over a period of time (longitudinal) or to collect data at one point in time. Surveys are designed to collect factual data related to sociological and psychological constructs such as opinions, attitudes, values, beliefs, and desires. Data are gathered in different ways, for example, with paper-and-pencil questionnaires, one-on-one interviews, focus groups, telephone interviews, electronic online questionnaires, and computerized kiosks. Responses are aggregated and summarized, and participant identity must be confidential or anonymous. Results of survey data are often reported by frequency and percentage for each response item. Other statistical procedures for reporting survey data are cross-tabulations (cross-tabs), chi-square statistic, phi coefficient, Kendall coefficient, and the gamma statistic.

Correlation

Correlation research is used to explore relationships between or among two or more variables. Correlation studies are useful for establishing predictive validity, establishing test reliability, and describing relationships. Simple correlation procedures involve ascertaining the relationship between two variables, whereas partial correlation procedures are used to control for a variable that may influence the correlation between two other variables. A multiple correlation coefficient (multiple regression) indicates the relationship between the best combination of independent variables and a single dependent variable. Canonical correlation indicates the relationship between a set of independent variables and a set of dependent variables. The kind of correlation coefficient computed depends on the type

of measurement scale used to collect the data and the number of variables.

Marie Kraska

See also Descriptive Statistics; Inferential Statistics; Standardized Tests; Statistical Significance; *T* Scores

Further Readings

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research.* Chicago: Rand McNally.

McMillan, J., & Schumacher, S. (2006). *Research in education: Evidence-based inquiry* (6th ed.). Boston: Allyn & Bacon.

Myers, R. H. (1990). *Classical and modern regression with applications* (2nd ed.). Pacific Grove, CA: Duxbury.

Ramsey, F. L., & Schafer, D. W. (2002). *The statistical sleuth* (2nd ed.). Pacific Grove, CA: Duxbury.

Salkind, N. J. (2003). *Statistics for people who think they hate statistics*. Thousand Oaks, CA: Sage.

Tukey, J. W. (1960). Conclusions vs. decisions. *Technometrics*, 2, 423–433.

Tukey, J. W. (1977). *Exploratory data analysis.* Reading, MA: Addison-Wesley.

R

He who lends a book is an idiot. He who returns the book is more of an idiot.

—Arabic Proverb

RANDOM SAMPLE

It is not always possible or feasible to collect data from every member of a population because of time constraints, lack of resources, or difficulty in identifying or locating all members of a specified group. In these cases, the researcher must select a sample from the larger group known as the target population to whom the researcher wants to generalize. Random samples allow the researcher to use inferential statistics to estimate the extent to which the sample differs from the population.

The target population is the group to which the researcher wishes to generalize, and the sample is a finite subset of the population. For example, if the researcher wished to generalize about the effect of comprehension strategies on first-grade students' reading ability, first-grade students would constitute the population. This group of first-grade students may be considered a sample of present and future first-grade students (sometimes referred to as a time/place sample) in one school, a school district, or state.

Usually one or more samples from a target population are used in a study. In these cases, researchers should follow systematic procedures for identifying the population, selecting the sample (subjects), and assigning subjects to treatment and control groups if findings are to be generalized beyond the study group

back to the target population. A representative sample helps to ensure that the generalizations or inferences made from the sample to the population are free of systematic bias and present a fair view of the target population. The best way to achieve a representative sample is through random sampling techniques. Different procedures may be used to select a random sample. Random sampling is often referred to as simple random sampling to distinguish it from other kinds of sampling, such as stratified random sampling, cluster sampling, and systematic sampling.

A *simple random sample* is the most basic kind of probability sample. Probability sample selection procedures are those in which each unit or subject in the population has a known, nonzero chance of being selected. For simple random sampling, each case in a finite population has an equal chance of being selected, and the selection of any one case is independent of the selection of any other case. For a sample to be truly random, systematic procedures should be employed to select the subjects. The recommended method for selecting a sample using simple random procedures is to use a table of random numbers. Tables of random numbers are usually computer generated, and they are often included in textbooks and on the Internet. To use a table of random numbers, the researcher first assigns an identification number to each subject in the population. Next, the researcher randomly selects a starting place on the table. This should be done by blindly placing a pointer on the

table to identify the row and column from which to start selecting numbers. Only those subjects whose identification number matches the number in the table are included in the sample. For example, to select a sample of 100 students from a list of 600, the researcher might number the students from 000 to 599, refer to the table, and systematically move through the table selecting only those students whose identification numbers fell between 000 and 599 until 100 students were selected.

Stratified random sampling procedures may be used when there are subgroups within a target population that the researcher wants to be sure to sample. For example, intelligence level, socioeconomic status, political affiliation, and number of children per household may form subgroups within a population. The strata or subgroups are first identified and then procedures are employed to select subjects randomly from each of the strata. The number of subjects selected for each stratum may be based on their proportional representation in the population, or equal numbers of subjects may be selected from each stratum.

Cluster sampling involves the use of intact or naturally occurring groups within a population. Individuals within the clusters should be similar in terms of the variables of interest to the researcher. Clusters should be selected randomly from a population of clusters, and all members of the clusters should be included in the sample. For example, in educational research, all students within certain grade levels, subjects, or classrooms are often used in a sample. Large samples such as those used by national education databases often use cluster sampling. However, in these cases, individuals within clusters are selected at random. For example, a representative sample of all students enrolled in the public schools in the United States may be drawn by random selection from each of the following clusters: public school districts in the nation, schools within the districts, classrooms from the selected schools, and students within classrooms.

Systematic sampling requires random selection of the first subject and then the selection of every *n*th subject from the target population list. The size of the interval to use to select subjects from the list can be calculated by dividing the population by the desired sample size. For example, the interval for a sample size of 500 from a population of 5,000 could be obtained by dividing the population size by the desired sample size (5,000/500 = 10). Every 10th subject would be selected for the sample. The starting point on the list should be identified randomly by rolling a pair of dice, using a table of random numbers, or using computer-generated numbers. If your random number were eight, then the eighth, eighteenth, twenty-eighth, and so on, subjects would be included in the sample. One should use caution when using systematic sampling to be sure that the population list is current and complete. Also, care should be taken that the population list is not formed by some distinguishing characteristic such as geographic location of subjects, which could introduce bias.

Probability samples are important for quantitative research methods because they permit statistical procedures whereby inferences can be made from a sample back to the target population. Regardless of the procedures used to select a sample, the most important characteristic of a sample is whether or not it is representative of the population to which the researcher wishes to generalize.

Marie Kraska

See also Quantitative Research Methods; Statistical Significance

Further Readings

Johnson, B., & Christensen, L. (2004). *Educational research: Quantitative, qualitative, and mixed approaches* (2nd ed.). Boston: Pearson Education.

Thompson, B. (2006). *Foundations of behavioral statistics: An insight-based approach.* New York: Guilford.

Tuckman, B. W. (1999). *Conducting educational research* (5th ed.). Belmont, CA: Wadsworth/Thomson.

READING COMPREHENSION STRATEGIES

A distinction is often made between skills and strategies. Skills are often conceptualized as unitary and automatic. Strategies are conceptualized as complex and effortful. Lysynchuk et al. defined comprehension strategies as "steps or actions that readers can take to enhance comprehension" (p. 460). Other available definitions are similar, but often have subtle distinctions. The National Institute for Literacy defines comprehension strategies as "conscious plans or sets of steps that good readers use to make sense of text."

The National Reading Panel of the National Institute for Child Health and Human Development defined comprehension strategies as "specific procedures that guide

students to become aware of how well they are comprehending as they attempt to read and write" (Sec. 4, p. 5).

What is important about these definitions is that they illustrate the diversity in thinking about strategies. However, the definitions all converge on the notion that strategies are cognitive, involving steps or actions, perhaps involving a number of skills. They also suggest that strategies are directed at enhancing, or making possible, the understanding of text.

The National Reading Panel reviewed the experimental research literature on comprehension strategy instruction and found a total of 203 research studies involving 16 different types of instruction of comprehension strategies. Of this total, 7 types of instruction of comprehension strategies were found to have significant scientific support in the literature reviewed by the panel. The 7 types are as follows:

1. *Comprehension monitoring* is a strategy in which readers learn how to be aware or conscious of understanding during reading. Readers also learn procedures to correct problems in comprehension as they arise.

2. *Cooperative learning* is a strategy in which readers work together to learn from reading. Students may work together on an entire problem or individually on components, sharing the results to complete the understanding.

3. *Graphic and semantic organizers* are strategies that allow readers to represent graphically (write or draw) the meanings and relationships of the ideas that underlie the words in the text.

4. *Story structure* is a strategy in which the reader learns to ask specific questions about the elements of stories, including plot, time line, characters, and events in the stories.

5. *Question answering* is the most conventional strategy in which the reader answers questions posed by the teacher (or the textbook) and is given feedback on the correctness of the answer.

6. *Question generation* is a strategy involving the active production of questions in which the reader asks himself or herself what, when, where, why, what will happen, how, and who questions.

7. *Summarization* involves the reader identifying and writing the main or most important ideas that integrate or unite the other ideas or meanings of the text into a coherent whole.

There were six additional categories of strategy instruction that did not have scientific support. These categories had too few studies or studies in which the range of student abilities or ages was too small to make judgments about scientific merit. This is an absence of knowledge rather than a judgment that these are ineffective. The categories were Curriculum (integrating strategy instruction into the normal curriculum), Listening Actively (oral language practice in comprehension in memory), Mental Imagery (picturing concepts mentally), Mnemonic (using memory devices), Psycholinguistic (instruction in language elements), and Vocabulary-Comprehension Relationship (emphasizing word knowledge).

Instructional Variables

Besides the content of instruction, there are several issues in how the instruction is presented or formatted. One important variable in the instruction of comprehension strategies is whether strategies are taught singly or in combinations. Reciprocal teaching is one of the more successful comprehension strategy instruction models. It incorporates four strategies: predicting, clarifying, questioning, and summarizing. Research has shown that the questioning and summarizing are dominant contributors to the success of the format.

Another variable in instruction of strategies is the relative amounts of student and teacher involvement. Two general models have developed.

In the Direct Explanation (DE) approach, teachers focus on helping students to conceptualize reading as a problem-solving task that requires strategic thinking. It also emphasizes the need to learn to think strategically about solving reading comprehension problems. A key element of this is the explicit explanation of the mental processes involved.

The Transactional Strategy Instruction (TSI) approach includes some of the same elements that DE includes. However, it also focuses on the ability of teachers to facilitate discussions in which students collaborate to form interpretations of text and explicitly discuss the mental processes and cognitive strategies that are involved in comprehension. The emphasis is on the interactive exchange among learners in the classroom.

The preparation of teachers to instruct students in comprehension strategies has also been studied extensively. These findings, also summarized by the National Reading Panel, show that teacher preparation to teach reading comprehension strategies is necessary

and can be accomplished, although it requires intensive and sustained effort.

Limitations of Current Knowledge

Instruction of comprehension strategies has been successful with students in the fourth grade and above. Very little is known about whether certain strategies are more appropriate for certain ages and abilities. Very little is known about the reader characteristics that influence successful instruction and which strategies, or combinations, are best for younger readers, poor or below-average readers, learning-disabled and dyslexic readers, or second-language learners.

It will be important to know whether successful instruction generalizes across different text genres (e.g., narrative and expository) and texts from different subject content areas. The National Reading Panel review of the research indicated that little or no attention was given to the kinds of text that were used and that there was little available information on the difficulty level of texts.

Michael L. Kamil

See also Instructional Objectives

Further Readings

Block, C. C., & Pressley, M. (Eds.). (2002). *Comprehension instruction.* New York: Guilford.

Lysynchuk, L. M., Pressley, M., D'Ailly, H., Smith, M., & Cake, H. (1989). A methodological analysis of experimental evaluations of comprehension strategy instruction. *Reading Research Quarterly, 24,* 458–470.

National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups.* Washington, DC: Author.

Web Sites

National Institute for Literacy: http://www.nifl.gov

RECIPROCAL DETERMINISM

Reciprocal determinism is a phrase coined originally by psychologist Albert Bandura in describing the workings of his social learning theory. It describes the reciprocal relationship between person variables (genetics, individual differences); overt behavior; and the social environment. Simply stated, it posits that a person's behavior is both influenced by and is influencing a person's personal factors and the environment. According to this view, a person's behavior can be conditioned by the environment through operant conditioning (e.g., the use of consequences like reward and punishment), but a person's behavior can also have an impact on the environment. Thus, it is not the case that individuals are merely shaped and influenced by their environment; individuals also influence the environment around them—each affects the other.

Theoretical Significance

Reciprocal determinism represented a significant departure from the long-standing debate in psychology over whether the person or the situation is most responsible for behavior. This new perspective argues that a person's behavior is based on his or her evaluation of the situation, and it added person variables and cognition to the factors that must be understood in order to have a full grasp of psychological phenomena. As such, this perspective influenced the way in which behavior is conceived, particularly within the realm of social psychology and social learning. It emphasizes the importance of person factors, such as cognition, in understanding the impact of the external environment on individuals as well as the ways in which personal factors affect the environment.

Specifically, this view discussed five different cognitive social learning person variables that affect an individual's evaluating and interpretation processes:

1. *Competency and self-efficacy.* Competencies include intellectual abilities, social and physical skills, and other abilities. Self-efficacy refers to the confidence an individual has in his or her ability to take action and persist in action. These variables address the question, "What can you do?"

2. *Encoding strategies and personal constructs.* People differ in the way they selectively attend to information, encode, or mentally represent events, and group information into meaningful categories. The same event may be perceived by one person as threatening, but by another person as challenging.

These variables address the question, "How do you see the situation?"

3. *Expectancies.* Expectations about the consequences of different behaviors guide an individual's choice of behavior. Before people engage in an action, they consider their beliefs about the likely results of that action. For instance, if you cheat on an exam and get caught, what do you expect the consequences to be? These variables address the question, "What will happen?"

4. *Subjective values.* Individuals who have similar expectancies may choose to behave differently because they assign different values to the outcomes. Two people may expect that donating money to charity will impress their peers; however, this outcome may be more important to one person than another, thereby affecting their choice of behavior. These variables address the question, "What is it worth?"

5. *Self-regulatory systems and plans.* People differ in the standards and rules they adopt for regulating their behavior (including self-imposed rewards for success or punishments for failure), as well as their ability to make realistic plans for reaching a goal. These variables address the question, "How can you achieve it?"

These person variables mediate people's interaction with the outside world. People's perceptions of their environment determine their actions, and the ways their social environment responds to them provides feedback that in turn affects these person variables. The reciprocal dynamics of this perspective recognize the degree to which various circumstances affect individuals, the way in which individuals interpret these circumstances, and the actions of the individuals themselves and others around them that serve to further change and define the situation.

Predicting and Understanding Behavior

In understanding behavior, reciprocal determinism emphasizes that an event and a response rarely have only a single cause. Furthermore, any events and responses that occur are likely to have an impact on future events as well. This perspective generates a complex view of causality, in which previous occurrences affect present occurrences, which in turn influence future events. This complex causality can best be demonstrated by an example.

Suppose a child named Billy approaches some other children on the playground and asks to join their game. The other children refuse to let Billy join because they are in the middle of a game. Because of Billy's own person variables, he interprets their response to be hostile toward him and retaliates by behaving aggressively toward his peers. In response, the other kids dislike Billy and start to avoid him and talk about him behind his back. Their response makes Billy even more paranoid about his peers' hostility toward him and reinforces his desire to aggress against them. Previously innocuous situations now become problematic because Billy no longer trusts his peers, and they in turn want to avoid him and exclude him from their games, leading to further acts of aggression. His persistent aggression toward his peers is observed by the teachers, who form a negative view of Billy and further reinforce his interpretation of the social environment.

This analysis illustrates the central tenets of reciprocal determinism—although the environment shapes, maintains, and constrains human behavior, people are not passive recipients in this process, for they actively choose, create, and change their environments. Thus, reciprocal determinism readily acknowledges that genes and other person factors affect behavior, and that environment strongly influences behavior as well. However, reciprocal determinism advocates the importance that subjective evaluation and interpretation play in determining behavior and in shaping and changing the environments people encounter. In utilizing reciprocal determinism as a tool for research and understanding, an emphasis is placed on dynamic changes over time that occur to all three factors (person variables, behavior, and environment) because of the complex interplay between these factors.

Edward Raymond Hirt and Timothy Scott Reilly

See also Applied Behavior Analysis; Behavior Modification; Social Learning Theory

Further Readings

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Upper Saddle River, NJ: Prentice Hall.
Mischel, W. (1973). Toward a cognitive social learning reconceptualization of personality. *Psychological Review, 80,* 252–283.

REGRESSION

Research in educational psychology is often interested in examining the relationship between one or more variables and using this relationship to predict future behavior. Regression is the process of modeling the relationship between variables. In statistical form, regression models can be used to predict the outcome of a particular variable given knowledge of the value of one or more other variables.

Regression can be used to answer questions related to performance on standardized examinations in subject areas such as mathematics, reading, and science. For instance, student performance on a mathematics assessment is likely related to performance in a mathematics-related class such as algebra or geometry. The relationship between these variables can be used to devise a regression equation that allows a prediction to be made. That is, given knowledge of performance in an algebra class, the regression equation can be used to predict (with error) performance on the mathematics assessment. The regression equation also explains a certain amount of variability in the outcome measure.

Regression models take many forms, but linear models are most common. Linear regression models refer to a pattern of change where the outcome measure or dependent variable (DV), Y, increases (decreases) as the independent variable (IV), X, increases (decreases). Theoretically, the linear regression equation can be expressed as

$$Y = a + bX + e,$$

where a and b are estimated values for the intercept and slope, respectively, and e represents measurement error or the amount of unexplained variability in the DV. The intercept is defined as the value of the DV when the IV is equal to zero, and the slope provides information on the direction and magnitude of the relationship between the IV and DV.

The goal in formulating regression models is to minimize the amount of measurement error by explaining as much variability in the DV as possible. This can be accomplished through either a simple regression model or a multiple regression model. The aforementioned example is considered a simple regression model using a single IV, performance in an algebra class, to predict performance on the mathematics

assessment. Including an additional variable, such as performance in a geometry class, would likely result in a better explanation of the variability in mathematics assessment scores. Theoretically, the multiple regression equation takes the same form as the simple regression equation but contains an additional slope value:

$$Y = a + b_1 X_1 + b_2 X_2 + e.$$

The slope in a simple linear regression equation is calculated as

$$b = \frac{s_{XY}}{s_X^2},$$

where s_{XY} is the covariance between X and Y and s_X^2 is the total amount of variability in X. The y intercept is

$$a = \overline{Y} - b\overline{X},$$

where $\overline{Y}$ is the mean of Y and $\overline{X}$ is the mean of X. The slope is reported in either standardized or unstandardized form. Unstandardized values of b explain the magnitude in the relationship between the independent variable and dependent variable in the original units of the variables. However, if the data are normalized, standardized values for the slope can be calculated. Using standardized values for the slope results in interpreting the relationship between X and Y in standard deviation units. Standardized values are useful for determining which IV is contributing more toward explaining the variance in the DV.

The IVs can also be compared in the amount of variability they are explaining via the coefficient of determination, also referred to as the squared correlation coefficient (r^2). The r^2 is a numerical value representing the amount of variability in the DV explained by the IV and is a measure of the goodness of fit of the regression model. For example, if data from two variables are graphed in a scatterplot, the coefficient of determination will be highest for the regression line that minimizes the distance between the observed data points and the regression line.

The coefficient of determination is calculated via the following formula:

$$r^2 = \frac{SS_{regression}}{SS_{total}} = \frac{\sum_{i=1}^{N} (\hat{Y}_i - \overline{Y})^2}{\sum_{i=1}^{N} (Y_i - \overline{Y})^2},$$

where $\hat{Y}$ is the predicted score of y from the regression equation. The $SS_{regression}$ (or sums of squares) represents the total amount of variability accounted for by the regression equation, whereas the SS_{total} represents the total amount of variability to be explained in the DV. Thus, the more variance explained by the regression equation, the larger the value of r^2. If a value of 1.0 was obtained, the observed data points would fall directly on the regression line, indicating a perfect fit between the model and the data. The coefficient of determination is interpreted as the proportion of variability in the DV accounted for by the IV. For example, suppose regressing performance in a mathematics class on performance on a standardized mathematics assessment yielded $r^2 = .25$. This indicates that approximately 25% of the variance in mathematics assessment scores is explained by performance in a mathematics class.

To determine the significance of the IV as a predictor of the DV, a test statistic is calculated. The test statistic uses information on the amount of variability accounted for by the model and the amount of variability left unaccounted for. The test statistic becomes larger as the model better explains the outcome measure. As more independent variables that contribute to explaining the variance in the independent variable are added to the model, the r^2 value will increase and the test statistic will also increase.

In summary, regression is tantamount to one of the principle purposes of behavioral research: to analyze and predict behavior. Linear regression can also be used for more complex research methods. For example, regression can be used for analyzing nonlinear relationships, analyzing the effects of covariates on a relationship, and determining the relationships between categorical independent variables (dummy coding). Regression also plays an important role in more advanced statistical techniques, such as structural equation modeling.

Greg W. Welch and Chris S. Meiers

See also Descriptive Statistics; Inferential Statistics; Statistical Significance

Further Readings

Pedhazur, E. J. (1997). *Multiple regression in behavioral research: Explanation and prediction* (3rd ed.). Fort Worth, TX: Harcourt Brace.

REINFORCEMENT

Reinforcement is defined as a consequence that increases the likelihood that a behavior will occur in the future. Understanding the use of reinforcement to teach new behaviors and strengthen existing behaviors is critical to the understanding of interactions between individuals. What may be a reinforcer at one time may not be the next time it is presented. A change in the power of a reinforcer is often based on many factors, such as satiation, deprivation, mood, age, or social interaction. Whether or not an event in one's environment is a reinforcer is determined by the effect it has on the behavior that precedes it, not on the intent of the person delivering it. An item is a reinforcer only if it increases the likelihood that the behavior will occur under similar circumstances in the future. Understanding that an item is only a reinforcer by its interaction with behavior is fundamental, and reinforcement, in this sense, can be related back to B. F. Skinner.

Positive and Negative Reinforcement

There are two types of reinforcement: positive and negative. Positive reinforcement is the addition of a stimulus that increases the likelihood that the behavior that preceded it will occur again. Negative reinforcement is the removal of a stimulus, such as escape from a task, which also increases the likelihood that a behavior will occur again. Negative reinforcement is often confused with punishment, but negative reinforcement increases the likelihood that the preceding behavior will occur again, whereas punishment decreases the likelihood.

Types of Reinforcers

There are functionally two types of reinforcers: primary (unconditioned) and secondary (conditioned). Primary reinforcers are events that individuals need and, by virtue of that, can be used as consequences to increase the likelihood that the preceding behavior will occur again. Examples include food, water, and sleep. Deprivation and satiation are especially important concepts when discussing primary reinforcers. A primary reinforcer that an individual has been deprived of will have extremely strong reinforcing

properties. Deprivation is limiting access to the item so that the individual strongly needs the item. An example of this might be if an individual has not drunk in several hours and, upon seeing water, might say the word "water." If the person is given the water, that water may act as a reinforcer, increasing the likelihood that saying the word "water" will happen again. In addition, if an individual satiates on an item that typically has reinforcing properties, that reinforcer may no longer have any reinforcing power.

Secondary reinforcers are events that become reinforcers after being paired with other reinforcing events. When delivered as a consequence, they also increase the likelihood that the preceding behavior will occur again. Examples might include stickers, social praise, or a toy. Often, a secondary reinforcer is delivered concurrently with a primary reinforcer or with another established secondary reinforcer. So, when first delivering a sticker as a reinforcer, a teacher might also deliver a small bite of food. With time and with systematic fading of delivery of the food, the sticker may have similar reinforcing properties as the food.

Another way that reinforcers are often distinguished is by type. Typically, primary reinforcers might fall in a category of edibles. Secondary reinforcers might include a variety of types, such as tangibles or activities.

Some examples of tangibles might include toys or stickers. Examples of an activity might be playing a game, going to lunch, or going for a walk. Social reinforcers might include tickling, high-fives, a smile, or verbal praise. These are things that are of a social nature, and the interaction is typically what develops as a reinforcer. Examples of generalized reinforcers might include tokens or money.

Schedules of Reinforcement

When looking at reinforcement, it is also important to look at the schedule of when reinforcers are delivered. Reinforcement can be continuous or intermittent.

Intermittent reinforcement schedules can be fixed ratio, variable ratio, fixed interval, or variable interval. Ratio refers to a certain number of responses being necessary before reinforcement is delivered. In a fixed ratio schedule of reinforcement, a set number of responses occur before the reinforcer is delivered. So, an FR6 (fixed ratio 6) schedule would mean that reinforcement would be given after six correct responses.

In variable ratio, the number of correct responses before reinforcement varies. A classic example of variable ratio is a slot machine.

Interval refers to a certain amount of time that must pass before a correct response will be reinforced. Similar to ratio, interval may be either fixed or variable. In fixed interval, reinforcement is delivered on the first correct response after a given interval, and in variable interval, the length varies.

Nanette L. Perrin

See also Applied Behavior Analysis; Behavior Modification; Classical Conditioning

Further Readings

Alberto, P. C., & Troutman, A. C. (1999). *Applied behavior analysis for teachers* (5th ed.). Columbus, OH: Merrill.

Baer, D. M., Wolf, M. M., & Risley, T. (1968). Current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis*, *1*, 91–97.

Baer, D. M., Wolf, M. M., & Risley, T. (1987). Some still-current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis*, *20*, 313–328.

Catania, A. C. (1997). *Learning* (4th ed.). Englewood Cliffs, NJ: Prentice Hall.

Cooper, J. O., Heron, T. E., & Heward, W. L. (1987). *Applied behavior analysis.* Columbus, OH: Merrill.

Iwata, B. A. (1987). Negative reinforcement in applied behavior analysis: An emerging technology. *Journal of Applied Behavior Analysis*, *20*, 361–378.

Michael, J. (1975). Positive and negative reinforcement, a distinction that is no longer necessary; or a better way to talk about bad things. *Behaviorism*, *3*, 33–38.

Skinner, B. F. (1953). *Science and human behavior.* New York: Free Press.

RELIABILITY

The two most important properties of an assessment are its validity and reliability. *Validity* refers to the meaningfulness of the interpretations and uses of a test score and is the most important property of an assessment. *Reliability* refers to the extent to which test scores are free from errors of measurement. Thus, validity examines the interpretations and uses that can reasonably be made from the consistent part of the test scores, whereas reliability is concerned with inconsistent or random errors of measurement. As a result, reliability is a necessary but not sufficient

condition for validity. That is, there needs to be some level of consistency to understand the meaningfulness of particular uses and interpretations of test scores, but measuring consistently does not guarantee the meaningfulness of the interpretations or uses.

Reliability and validity are not global properties of an assessment. Instead, they are properties of specific uses and interpretations that are made from a set of test scores. A test could be valid for a particular use or interpretation and not for another. For example, a test might measure the curriculum covered in a school without providing valid estimates of student performance because of the length of the tests or the nonequivalence of forms. The same is true for reliability. For example, a test might provide reliable scoring without being stable over time. In addition, reliability and validity are a matter of degree. Tests are not considered valid or invalid. Instead, they are valid to some degree. Similarly, a test is not considered reliable or unreliable, but is reliable to some degree.

Estimates of reliability are indices that quantify the amount of measurement error for a particular test use or interpretation for a specified population. Although reliability can be defined broadly in terms of consistency or generalizability, specific statistical indices of reliability will vary depending on the statistical model and the sources of error. The statistical model may be based on classical test theory, generalizability theory, or item response theory. Classical test theory and generalizability theory are based on total scores, whereas item response theory is based on an estimate of a latent trait. In this entry, only classical test theory and generalizability theory are considered. Within each theory, there are multiple indices of reliability based on multiple sources of measurement error, including item heterogeneity, equivalence of test forms, stability over time, and consistency of subjective ratings. Different sources of error would be of concern in different contexts. For example, the test score of a student writing an essay is affected by errors in scoring, whereas the test score from a student taking a multiple-choice test is affected by the heterogeneity of the items selected to measure the construct. In addition, a test score can be affected by multiple sources of error simultaneously. A student taking the GRE might be affected by the heterogeneity of the items, the form of the test, and the subjectivity of the scoring for the written portion of the test. Thus, there are many types of reliability that vary depending

on the sources of error being considered as well as the statistical model or test theory being used. These varying definitions will be selected based on the particular test use or score interpretation being made, and one type of reliability should not be considered interchangeable with another.

Classical Test Theory and Estimates of Reliability

Classical test theory assumes that any observed test score, X, is the sum of a true score, T, and a random error, E. That is, $X = T + E$. The issue of defining the true score is a matter of validity, whereas the issue of defining the random error is a matter of reliability. According to classical test theory, the error is the sum of all random components, whereas the true score is the sum of all consistent effects. Thus, the error is undifferentiated with respect to different sources of randomness, unlike in generalizability theory. Broadly, two indices of reliability are commonly reported: the reliability coefficient and the standard error of measurement.

The reliability coefficient (ρ) is defined as the ratio of the true score variance to the observed score variance, or the ratio of the true score variance to the sum of the true score variance and the error variance. Hence, the value of the reliability coefficient is the proportion of variation in test scores that can be attributed to consistent measurement (i.e., the true score). The reliability coefficient ranges from 0.0 to 1.0, with higher values being preferred. At $\rho = 0.0$, there is no consistency in the measurement procedure and the observed score is equal to random error ($X = E$). At $\rho = 1.0$, the observed score has no error and is equal to the true score ($X = T$). In practice, the reliability coefficient will be somewhere between the two extreme values.

The standard error of measurement (SEM) is the standard deviation of the errors of measurement. The SEM ranges from 0.0 to the standard deviation of the observed scores, σ_x. When the SEM $= \sigma_x$, there is no consistency in the measurement procedures, the reliability coefficient is equal to 0.0, and the observed score is equal to the random error. When SEM $= 0.0$, there is perfect consistency in the test scores, the observed score is equal to the true score, and the reliability coefficient is equal to 1.0. In practice, the SEM will fall somewhere between the two extreme values.

The reliability coefficient is an easily interpreted index of the consistency of the test scores because it is in a standard range for all tests. Although the SEM is more difficult to interpret initially, it is in the metric of the test scores that allows for the interpretation of the individual test scores via confidence intervals. Another advantage of the SEM is that it is not based on the true scores, and consequently, it is not influenced by sampling errors. The reliability coefficient will be underestimated when the sample range of scores is restricted, whereas the SEM will be largely uninfluenced by sampling fluctuations.

Types of Reliability

Within the framework of classical test theory, there are several types of reliability coefficients based on the source of the random errors. The types of reliability discussed below are test-retest, alternate form, alternate form test-retest, interrater, split half, and internal consistency.

Test-retest reliability is used to examine the stability of the trait being measured over time. The reliability coefficient is the correlation between test scores for a sample taking the same test on two occasions. Generally, test-retest reliability is higher when the time span between test administrations is shorter. However, the test-retest reliability should be estimated with a time interval that mirrors the actual use of the test rather than trying to maximize the value of the coefficient.

Alternate form reliability is used to measure the equivalence of test scores across two (parallel) forms of a test. The reliability coefficient is the correlation between test scores on the two forms of the test taken by the same sample without a substantial time lag. Usually, half of the sample receives one form first (e.g., Form A), and the other half of the sample receives the other form first (e.g., Form B) so that there is no order effect. Then, examinees take the form they have not taken yet. Alternate form reliability is higher when care is taken to make sure that the two forms are equivalent in content and statistical properties (i.e., mean, standard deviation, and distribution shape).

Alternate form test-retest reliability follows the same procedure as with the alternate form reliability except that there is a time lag between test administrations. In this case, the errors of measurement include stability over time and equivalence of the forms. In general, this type of reliability will be lower than alternate form or test-retest reliability, which target only one type of random error.

Interrater reliability is used to measure the consistency of ratings from subjective scoring. The reliability coefficient is the correlation between the ratings from two raters on the same sample of writings/essays. Interrater reliability is higher when standardized procedures are used by the raters to score the writings. At a minimum, the standardized procedures should include training of the raters and clearly defined rubrics. Large-scale assessments further standardize the procedures to include benchmark writings, monitoring the process, intervening when ratings disagree, and other procedures to check the rating process.

Split half reliability is used to measure the consistency within a single administration of a test by examining the relationship between two halves of the same test. The procedure for split half reliability is to administer a single form of the test to a sample. The reliability estimate is then based on the correlation between two halves of the test adjusted for test length. That is, the test is divided into two equivalent halves based on test content and item statistics (often, this can be accomplished by using odd- and even-numbered items to form the halves), and the halves are correlated. However, the reliability will be less for half of a test than it is for the full-length test. Consequently, the correlation between the halves is adjusted upward using the Spearman-Brown prophecy formula. Split half reliability will be higher when the equivalence of the two forms is higher in terms of content and item statistics. However, the matching of the two halves should not be completed on the basis of the sample statistics because random sampling fluctuations could inflate the value of the reliability. Instead, careful matching should be completed based on content and item statistics from a prior data collection.

Internal consistency is used to measure the consistency of items within a single test form. The procedure for internal consistency is to administer a single form of the test to a sample and estimate the internal consistency using item and test statistics with an internal consistency formula. The formula for internal consistency has many equivalent forms in the literature, including the Kuder-Richardson 20 (KR20) formula for dichotomously scored items and Cronbach's alpha. Internal consistency is also easy to compute with most standard statistical software (e.g., SPSS or SAS).

Internal consistency is higher when the items are more homogeneous.

Below is a summary of the reliability coefficients and their major sources of error that are reported in classical test theory:

Reliability Type	Source of Error
Test-retest	Stability over time
Alternate form	Equivalence across forms
Alternate form test-retest	Stability over time and equivalence across forms
Interrater	Consistency of ratings
Split half	Equivalence across halves
Internal consistency	Equivalence and item homogeneity

Each of the reliability coefficients above differs in its data collection procedure, computation, and major source of error. The "appropriate" reliability coefficient should match the intended use or interpretation of the test. For example, when subjective measurements are part of the assessment procedure, interrater reliability is needed. When multiple items are being used (which should be the case), internal consistency or split half reliability should be used. In short, the reliability estimate(s) should include all sources of error that will be part of the test use or interpretation. One type of reliability should not substitute for another.

Standard Error of Measurement

The reliability coefficient is used to quantify the precision of an assessment for a particular use or interpretation. The index is simple to interpret because it is always based on the same scale (0.0–1.0). However, the reliability coefficient fails to show the amount of error that might be expected in an individual's test score. The SEM is the standard deviation of the errors of measurement and can be used to create confidence intervals for examinee scores. Assuming a normal distribution, 68% of the observed scores will be within one SEM of their true score, and 95% of the observed scores will be within 1.96 SEMs of their true score. For example, if SEM = 2.00 and an examinee's true score was 25, upon repeated

measurements, 68% of the scores for that examinee would be between 23 and 27. Note that the confidence interval is around the true score and not the observed score, which leads to the interpretation that 68% of the time that a confidence interval based on one SEM is constructed, it will contain the true score.

As with the reliability coefficient, a SEM can be created for different types of measurement error. In fact, the SEM is calculated using the appropriate reliability coefficient so that the appropriate source of error is being used. Thus, the table mentioned earlier can be used for the SEM or the reliability coefficient so that the SEM can be created with each of the different sources of error.

Magnitude of Reliability

The literature does not provide definitive guidance on acceptable levels of reliability. However, it is clear that what constitutes an acceptable level of reliability is determined by the use of the test. Uses of the test with higher stakes require higher levels of reliability. For example, reliability for a test being used in theoretical research may not require the same level of consistency as would be required for high-stakes uses of tests such as high school graduation, certification, or licensure.

How to Increase Reliability

It is important to be able to increase reliability when developing instruments. In general, there are two ways that should always be considered when increasing reliability: greater standardization and increasing the number of items. Test administration and test development procedures should be standardized so that no random errors are introduced. The effect of the standardization will not only globally affect each type of reliability, but it will also have specific effects on certain types of reliability. Standardization includes methods to create equivalent forms of a test (alternate form), methods to create homogeneous pools of items (internal consistency), or equivalent halves of tests (split half). Standardization also includes methods to create consistency in scoring through the development of rubrics and standardized scoring procedures (interrater).

Another key element to increasing reliability is increasing the length of the test. The Spearman-Brown formula is based on the principle that longer tests are

more reliable. Assuming that the conditions of testing do not change with increased length (i.e., fatigue, boredom, or item quality), increasing the number of items always leads to more reliable tests. Thus, short forms of a test or subscores are generally less reliable than the full form. As a consequence, subscores and short forms are also more difficult to interpret. Thus, a test that provides a reliable total score for accountability may not be useful when examining the subscores that might be needed for diagnostic interpretations.

Reliability and Aggregation

Increasing the number of items will increase the reliability of test scores because the scores are averaged over more data. Similarly, increasing the number of examinees and averaging across their scores will reduce the SEM and increase reliability. That is, the reliability of the mean will be higher than the reliability of an examinee. This applies to estimates for the full sample as well as aggregates such as classrooms, schools, or states. Whether averaging across items or examinees, the estimate becomes more stable. Indeed, the SEM will almost always be lower for group means than it is for individual means. (Note: under some conditions, the reliability coefficient could be lower for the group means when the true score variance in the groups is restricted in range. However, even under these conditions, the SEM will typically be lower and the group means will be more stable.)

Reliability and Growth Scores

The reliability of growth or difference scores, defined as posttest minus pretest, has received considerable attention in the literature. Some have argued that the growth score is unreliable and that growth is negatively correlated with the pretest. That is, growth will be higher for examinees with low pretests. However, other researchers have pointed out that low reliability for the difference scores does not necessarily result in less power for comparisons among groups, and the difference score may be the construct of interest. At any rate, caution should be used in interpreting growth scores at the examinee level.

Relationship of Reliability and Validity

Classical test theory assumes that an examinee's test score is the composite of a true score and random error. Validity addresses the true score by examining its uses and interpretations. Thus, any systematic error or bias is part of the true score, whereas only random errors are addressed in the reliability analysis. As discussed above, reliability is a necessary but not sufficient condition for validity. This means that there needs to be a true score (with some level of reliability) to examine validity, but that the existence of a true score does not guarantee that it is not a biased estimate of the construct of interest as a result of some systematic error.

In addition to this relationship of validity and reliability, there may be a tension between the two psychometric properties of the test. Higher reliability can be attained by standardizing the testing procedure, which has the potential to reduce the breadth of the construct being measured and, thus, to decrease the validity. For example, higher internal consistency is attained by increasing item homogeneity. To the extent that the construct requires heterogeneity of the items, this will create a tension between reliability and validity. As a second example, interrater reliability is increased by standardizing the scoring procedure (e.g., clearly defined rubrics and training). This standardization can limit the definition of good writing by not including some types of writing in the rubric and thus, as a consequence, limit the breadth of the construct. As a result, it is important to consider the impact of any standardization on the validity of the test as well as the reliability so that the construct is still clearly being measured.

Generalizability Theory and Estimates of Reliability

Classical test theory examines errors of measurement with a single undifferentiated error that may represent multiple sources of error (e.g., alternate form test-retest reliability). In addition, classical test theory examines reliability only from a norm-referenced perspective. That is, the methods rely solely on correlations that focus on rank ordering of scores. The correlations are sensitive to differences in rank ordering but not to shifts in scale. Thus, the reliability can be high even when the scales for the two forms, raters, and so on, are substantially different. For example, Rater A could rate 10 points higher than Rater B, and the reliability would be equal to 1.0 as long as every essay was scored exactly 10 points higher by Rater B than Rater A.

Generalizability theory solves each of these issues by (a) modeling multiple sources of error and (b) differentiating between errors based on rank ordering (i.e., relative error) and errors based on point estimation (i.e., absolute error). Generalizability theory assumes that each examinee has a universe score that is his or her average score across all conditions in the universe of admissible observations. That universe is composed of measurement facets (e.g., raters, items) with a particular level of a facet being a condition (e.g., selected raters or items). The potential measurement conditions selected from the study are then considered to be the universe of generalization. Generalizability theory is more complex statistically than classical test theory and is done in two stages: the G-study and the D-study. In the G-study (generalizability study), random effects analysis of variance (ANOVA) is used to estimate variance components for each of the effects in the model. The ANOVA, with the associated variance components, can be estimated with one or more facets (e.g., persons by items, or persons by items by raters). In the D-study (decision study), alternative measurement models can be examined to optimize the measurement procedures or to examine a reasonable set of measurement conditions. The results of the D-study identify the universe of generalization.

Similar to classical test theory, there are two types of indices computed in the D-study that show the consistency of the measurement procedure. The generalizability coefficient, or dependability index, is the ratio of the universe score variance to the universe score variance plus the error variance, and it is analogous to a reliability coefficient, whereas the second index is analogous to the SEM. The generalizability coefficient shows the ratio when using relative error variance and emphasizes the rank ordering of scores. Thus, it would be used for norm-referenced score reporting. The dependability index shows the ratio when using absolute error variance and emphasizes the absolute magnitude of the test scores. Thus, it would be used for criterion-referenced score reporting. The second index is the standard error. The standard error also can be computed for relative or absolute score reporting. The use and interpretations of the indices in generalizability theory are analogous to the indices in classical test theory.

Absolute error is always greater than or equal to relative error. Consequently, the generalizability coefficient is less than or equal to the dependability index, and the absolute standard error is greater than or equal to the relative standard error. As a result, more conditions (items, raters, etc.) may be needed when estimating examinee scores absolutely rather than relative standing.

Consistency of Classification

Each of the reliability coefficients, whether from classical test theory or generalizability theory, is based on continuous variables. Often, the measurement procedure is based on the classification of examinees. For example, the National Assessment of Educational Progress (NAEP) classifies students as Advanced, Proficient, Basic, and Below Basic. Clearly, when the data are nominal or categorical, a different statistical procedure must be used to examine consistency. Decision consistency is a method of examining reliability and the exact agreement across measurement conditions when the data are categorical. Decision consistency can be calculated for each of the sources of error described above in the section on classical test theory.

Two indices commonly reported for decision consistency are the proportion agreement and Cohen's Kappa. Proportion agreement shows the proportion of the examinees that are classified the same across forms, time, and so on. For example, if two forms were given to a sample for the NAEP reading assessment in Grade 8, the proportion agreement would be equal to the sum of the proportions that were Advanced on both forms, Proficient on both forms, Basic on both forms, and Below Basic on both forms. The same calculations could be done with other sources of error (e.g., ratings are the same across two raters). Cohen's Kappa is the proportion agreement after statistically adjusting for the expected agreement. Thus, Cohen's Kappa shows the agreement above and beyond chance and generally has lower values than the proportion agreement.

M. David Miller

See also Assessment; Descriptive Statistics; Evaluation; Testing

Further Readings

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing.* Washington, DC: American Educational Research Association.

Brennan, R. L. (2001). *Generalizability theory.* New York: Springer-Verlag.

Haertel, E. H. (2006). Reliability. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 65–110). Westport, CT: Praeger.

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores.* Reading, MA: Addison-Wesley.

Risk Factors and Development

Optimal child development may be compromised by several risk factors, including poverty, poor nutrition, harsh/inconsistent parenting, familial substance abuse, and low parental education, to name a few. Children developing within these maladaptive contexts often exhibit poor developmental outcomes across cognitive, social, and behavioral domains, such as low academic achievement and antisocial behavior. Some children exposed to great adversity, however, develop competence, suggesting that some factors may protect (or buffer) high-risk children from experiencing negative developmental outcomes. The purpose of this entry is to (a) describe the theoretical framework guiding the research on the development of risk and protective factors; (b) identify how risk factors at the child, parent, family, and community level negatively affect developmental outcomes; and (c) describe how protective factors may moderate negative outcomes in high-risk children. This entry also discusses the importance, capacity, and effectiveness of early intervention programs that are designed to increase educational outcomes by reducing the negative consequences of risk factors.

Ecological Systems Theory as a Guiding Theoretical Framework

Ecological frameworks recognize that each person functions within a complex network of individual, family, community, and environmental contexts that influence the availability of risk and opportunities to avoid risk. Urie Bronfenbrenner is a pioneer of the ecological model of human development, in which he describes development as the composite of individual genetic endowment, immediate family influences, and other components of the environmental context. According to Bronfenbrenner, the individual is embedded in five interrelated, nested subsystems that simultaneously influence the process of human development.

The innermost circle represents the microsystem of the individual. Within the microsystem are the individual's interactions with his or her immediate settings. Bronfenbrenner refers to these interactions as *proximal processes.* For most children, the family is the first and most important microsystem. As development proceeds through childhood and adolescence, additional microsystems might be sports teams, youth or church organizations, and work. Microsystems may overlap in that the same person may be a member of more than one system in a child's life. For example, a friend may be a member of the child's peer group, sports team, and class in school.

Just as the individuals interact with others in their microsystem, separate microsystems interact with each other at the level of the mesosystem. The mesosystem incorporates linkages between settings such as family, peers, teachers, and other school personnel. For example, an adolescent's ability to excel in school may depend more on the interconnections between the school and the home rather than solely on adequate performance in the classroom. In this case, Bronfenbrenner posits that the breakdown of connections between family, school, peer group, and neighborhood underlie the decline of academic achievement more so than relationships within each of these contexts alone.

Continuing the progression to increasingly distal influences on the individual, the exosystem represents external environments that the individual may or may not experience; yet events that occur in these environments affect what happens in the microsystem (an individual's immediate setting). This system includes features of the community such as availability of services or employment, access to formal and informal support, and socioeconomic climate. Parental and teacher social networks are part of the exosystem, having the potential to influence interactions with the child even though they may not directly be experienced by that child.

The overarching macrosystem represents the larger societal contexts, and influences differ across socioeconomic, ethnic, religious, and other subcultural groups. For example, in the United States, macrosystem influences include American culture and the social policies and programs that affect American families, such as racial prejudice and discrimination, the media, and antipoverty social programs. Finally,

each subsystem is embedded with time, the chronosystem, which represents that each of these subsystems functions within time and changes over time.

Applying this model to one of the most researched child outcomes, low academic achievement, indicates that risks begin in the most proximal system. The microsystem includes risk factors such as impulsivity, poor attention, inadequate school readiness, poor physical health, low school attendance, and so on. In the mesosystem, low academic achievement relates to the interactions between the family and the school, and include such factors as parenting styles, parent involvement in academic endeavors, teaching styles, school environment, and so on. The community context also affects a child's ability to perform well academically, given such factors as community violence, neighborhood crime, parent employment, school policies, transportation availability, and so on. Finally, through cultural values, governmental policies, and the political climate, the larger society also begins to influence academic achievement. Given the influences from multiple levels of children's ecology, it becomes evident that the risks for low academic achievement cannot be understood from one context alone.

Risk Factors Across the Ecological System

The structure of the ecological model allows for understanding the influences on child development at multiple levels and informs the process by which contextual factors influence child development. Although this framework identifies risk factors for child development, the concept of risk initially evolved from medical research examining the cause of cardiovascular disease. A medical research study that began in 1948, called the Framingham Heart Study, followed multiple cohorts of individuals across several decades in an effort to identify the factors that put individuals at risk for developing cardiovascular disease. The same logic has been applied to human development and enhanced understanding of why some children facing adversities from multiple ecological levels develop into cognitively and socially competent adults, and others do not. Several longitudinal studies have documented risk factors and the decline of optimal child and adolescent development (e.g., the Perry Preschool Project, the Abecedarian Project, the Milwaukee Project, and Project CARE).

As research studies began to investigate the cause of poor developmental outcomes and risky behavior (e.g., low academic achievement, teen pregnancy, juvenile delinquency), three categories emerged: child characteristics, parent/family characteristics, and community characteristics. Although risk factors are often studied and discussed within these categories, the ecological framework informs the interrelated nature of each factor and ecological system.

Child Characteristics

The three main risk factors associated with poor developmental outcomes include low birth weight, temperament, and child health. First, compared to infants born at normal weight (> 5 lbs.), children born with low birth weight (1 to 4 lbs.) are more likely to have continual health and developmental problems (e.g., cerebral palsy); lower IQ; and pervasive neuropsychological impairments in attention, memory, problem solving, reasoning, and language. Marked negative behavioral and academic outcomes are also evident, including higher incidence of conduct disorder, hyperactivity, attention problems, and learning disorders. Furthermore, the developmental consequences of low-birth-weight babies are particularly devastating for the extremely low-birth-weight babies (less than 2 lbs.).

The second risk factor at the child level is temperament. Children with difficult temperaments are at much higher risk for poor developmental outcomes than easy-temperament babies. Difficult children have a negative effect on their caregivers and thus are less likely to experience supportive relationships with them, which is a necessary component to counteract exposure to adversity. Difficult babies often develop into difficult adolescents, who have been shown to exhibit more emotional and behavior problems than adolescents who were easy babies.

The third risk factor associated with poor developmental outcomes is poor child health. Frequent hospitalizations, chronic medical conditions, and repeated illnesses increase the stress level within the family, which has been related to later psychosocial and academic difficulties in school-age children. Poor health in childhood has a long-term impact on development across the life span and is related to early mortality rates.

Parent/Family Characteristics

Research has uncovered several parental and family-level risk factors associated with maladaptive

development, including family structure and resources, parental background characteristics, and relationships within the family. One of the most pervasive negative effects on developmental outcomes is a family living in poverty. Research suggests that children raised in persistent poverty (as opposed to transitory poverty) perform much worse academically—as measured by achievement test scores, grade promotion, high school graduation, and completed years of schooling—than more affluent children. Furthermore, children in economically strained environments are almost twice as likely as nonpoor children to experience learning disabilities and developmental delays. Parent, teacher, and child self-reports also reveal that low-income children display more externalizing problems, such as fighting, difficulty getting along with others, and impulsivity, in comparison to higher-income children. Finally, researchers report that children of poor families have more emotional and behavioral problems than do children who have never been poor.

Vonnie McLoyd reviewed the research on children of low-income families. Results indicate that family income produces the strongest relationship with academic achievement of all other variables considered (e.g., parent's occupation and parent's education level). Even when the other indicators are summed together, they do not yield as high a correlation with academic achievement as does income alone. This line of research shows that economic hardship influences the psychological well-being of parents and the relationships among family members. Studies reveal that the influence of economic stress on children and adolescents is mediated to a significant extent by disruptions in family relationships and interactions, especially parent–child interactions.

In addition, families headed by single parents are also six times more likely to be poor than are families from two-parent households and therefore are more likely to be affected by poverty characteristics. Single parenthood disrupts the relationship between children and the noncustodial parent, usually the father. This may affect the extent to which supportive, caring relationships can be developed and nurtured. Psychological problems and behavior problems are more likely in children from single-parent or remarried families.

Another risk factor can be conceptualized as parental background characteristics, which includes parental substance abuse and low parent education. Each has been linked to increased internalizing and externalizing problems among children. Also, children

with parents suffering from alcoholism are approximately four to six times as likely as the general population to develop alcohol problems. Furthermore, anxiety, depression, and externalizing behavior disorders (e.g., conduct disorder) are more common among children with alcoholic parents than among children of nonalcoholics. In families where alcohol or other drugs are being abused, behavior is frequently unpredictable, communication is unclear, and family structure is either nonexistent or inconsistent. Because parents who abuse alcohol or other drugs are more likely to be involved with domestic violence, divorce, unemployment, mental illness, and legal problems, their ability to parent effectively is severely compromised. There is a higher prevalence of depression, anxiety, eating disorders, and suicide attempts among children of parents who abuse drugs and alcohol than among their peers.

Family background factors, such as parental education, have moderate effects on child development but larger effects on later adult outcomes. That is, parental education level affects later adult outcomes through transmission of educative values, such as not dropping out of school. Parent income and education level are associated with school drop-out rates, which drastically reduce the quality of later adult life, increasing the likelihood of unemployment and poverty.

The final risk factor to be discussed here includes the quality of the relationships within the family (i.e., parent–child, sibling, and marital relationships). In general, research reports that warm and supportive relationships within the family are associated with positive child and adolescent outcomes, whereas coercive and conflictual relationships are associated with the development of problems. One of the clearest and most replicated associations in this area is between conflictual parent–child relationships and child and adolescent maladjustment. For example, harsh, inconsistent, and ineffective discipline is associated with later antisocial behavior, and insecure parent–child relationships, including parental rejection and forcefulness, are traced to later internalizing problems.

Community Characteristics

The community context can also influence a child's developmental trajectory. Research has shown that poor neighborhoods can have detrimental effects on individual health status through three types of pathways. First, concentration of poverty and related characteristics may

create more detrimental social environments (e.g., violence, stress and anxiety, exposure to drugs, limited social control). Second, poorer communities are less likely to have access to adequate health care and social services. Third, the physical environment (e.g., air pollutants, hazardous conditions leading to accidents, poorer sanitation) in poor communities may be worse than in more affluent communities.

Existing research also points to a powerful connection between residing in an adverse environment and participating in criminal acts. Inner-city neighborhoods provide limited economic, institutional, and social resources for families and adolescents. Disorganized neighborhoods have weak social control networks, which result in isolation among residents, high residential turnover, and criminal activity that goes unmonitored. Few opportunities for youth to be monitored (such as programs where youth can contribute to the community), as well as the increased access to alcohol, drugs, and firearms, increase the likelihood of youth engaging in risky and destructive behaviors. Neighborhoods with high crime, poverty, and physical disorder are associated with youth engagement in violence, crime, and drugs and alcohol; low community attachment; and a feeling of being unsafe.

The Impact of Cumulative Risk

Although each risk factor across the ecological system has been related to a negative cognitive, social, or behavioral impairment, the impact of one risk factor tends to be relatively small in relation to the cumulative effect of multiple risk factors. That is, research suggests that children can handle some level of adversity, but when faced with three to four risk factors, developmental outcomes become drastically and negatively affected. Furthermore, the more risks to which children are exposed, the worse the developmental outcomes. Therefore, children's optimal development is ultimately affected by the sheer number of risk factors present, regardless of the nature of a particular risk factor. As the number of risk factors increases, the negative effect enlarges disproportionately. For example, having four or more risk factors, which relate to the child, parent, and sociodemographic situation, can lead to a 10-fold increase in difficulties, a result that has been replicated numerous times across the psychological research.

It is important to note that although poor developmental outcomes tend to relate to the presence of risk factors, the absence of risk factors does not predict healthy development. Appropriate opportunities must be presented for development to flourish and reach optimal levels. For example, opportunities afforded by high parental warmth, rather than just a lack of harsh and inconsistent parenting, are essential to promote healthy development. Research indicates that development is a balance between opportunities for strengths and risks to develop. Within each of the ecological levels, there are risk factors that negatively influence development and there are protective factors that foster positive development. Risks, cumulative risks, and protective factors interacts generating a complex set of relationships that creates development. Knowledge of risks alone is not sufficient to understanding child development, although these factors have received the most research attention.

Modifying the Impact of Risk Factors on Development

The negative impact of cumulative risks may be offset by understanding the conditions that promote successful development. Resiliency research has demonstrated that there are certain factors that protect against undesirable behavior. Resiliency has been defined as the capacity of the person, family, or community to prevent, minimize, overcome, or thrive in spite of negative or challenging circumstances. There are three kinds of resiliency described in the research literature. The first involves an individual who overcomes all the odds, suggesting that this individual has a particular quality that allows him or her to withstand adversity. The second consists of coping in the face of sustained and acute negative circumstances, such as extremely high family conflict. Third, resilience can refer to recovery from trauma, such as the death of a sibling. All require risk to be present for resilience to emerge; however, most research tends to fall within the second group: success in the face of adversity.

Identifying protective factors is one way that resiliency can be measured. Protective factors are those that relate to positive developmental outcomes for youth exposed to high levels of risk or adversity only. If positive outcomes are observed across all levels of risk (high and low), the factor is termed an asset, compensatory, or promotive. Thus, protective factors are assets that particularly matter when risk or adversity is *high*. Protective factors *moderate* the impact of

adversity on adaptation, changing the expected negative child outcome.

A landmark study on resiliency, conducted by Emmy Werner, has followed the development of children born on the Hawaiian island of Kauai since 1955. It has provided a wealth of data on protective factors for positive development in children with high cumulative risk. In this study, the risk group (about one third of the children) was defined by having four or more early risk factors that included poverty, perinatal stress, family conflict, and low parental education. About one third of these high-risk children developed well in terms of getting along with parents and peers, doing fine in school, avoiding serious trouble, and having good mental health. The resilient group had more resources and fewer adversities from an early age, including (a) close attachment to at least one good parent, (b) family harmony (i.e., conflict-free environment), (c) lack of parental psychopathology, (d) more time before the next child in the family came along, (e) easier temperaments as babies, (f) better intellectual skills and competence, (g) more connections with prosocial adults, (h) fewer separations from caregivers, and (i) better overall physical health. Children were also responsible, self-confident, and motivated to achieve.

On the other hand, the nonresilient group exhibited many of the same negative developmental outcomes associated with high-risk youth described earlier. The follow-up study followed high-risk children who were exposed to chronic poverty, birth complications, parental psychopathology, and family discord into adulthood. With the exception of serious central nervous system damage, the impact of adversities during childhood diminished adult adaptation depending on the quality of the child-rearing environment and the emotional support provided by family members, friends, teachers, and adult mentors. Poorest outcomes at age 40 were associated with prolonged exposure to parental alcoholism and/or mental illness. Men and women who had encountered more stressful life events in childhood reported more health problems at age 40 than did those who had encountered fewer losses and less disruption in their family during the first decade of life. This study demonstrates the need for early attention to the health status of our nation's children, especially those who are exposed to poverty, serious perinatal complications, and parental psychopathology. The policy implications are clear: Early access to good preventive and ameliorative health services and proper attention to the quality of early child care can result in improved quality of life in adulthood.

Implications for Educational Psychology

As a group, children who live in poverty tend to perform worse in school than do children from more privileged backgrounds. For the first half of the 20th century, researchers attributed this difference to inherent cognitive deficits. At the time, the prevailing belief was that the course of child development was dictated by biology and maturation. These preliminary results caught the attention of Sargent Shriver, President Lyndon Johnson's chief strategist in implementing an arsenal of antipoverty programs as part of the War on Poverty. His idea for a school readiness program for poor children focused on breaking the cycle of poverty. Shriver reasoned that if poor children could begin school on an equal footing with wealthier classmates, they would have a better chance of succeeding in school and avoiding poverty in adulthood. He appointed a planning committee of 13 professionals in physical and mental health, early education, social work, and developmental psychology. Their work helped shape what is now known as the federal Head Start program.

The three developmental psychologists in the group were Bronfenbrenner, Mamie Clark, and Edward Zigler. Bronfenbrenner convinced the other members that intervention would be most effective if it involved not just the child but the family and community that comprise the child-rearing environment. Parent involvement in school operations and administration was unheard of at the time, but it became a cornerstone of Head Start and proved to be a major contributor to its success. Zigler had been trained as a scientist and was distressed that the new program was not going to be field-tested before its nationwide launch. Arguing that it was not wise to base such a massive, innovative program on good ideas and concepts but little empirical evidence, he insisted that research and evaluation be part of Head Start. Although it is difficult to summarize the hundreds of empirical studies of Head Start outcomes, high-quality Head Start programs do produce a variety of benefits for most children. Although some studies have suggested that the intellectual advantages gained from participation in Head Start gradually disappear as children progress through

elementary school, some of these same studies have shown more lasting benefits in the areas of school achievement and adjustment.

Head Start began as a great experiment that has yielded prolific results over the years. Some 20 million children and families have participated in Head Start since the summer of 1965; current enrollment approaches one million annually, including those in the new Early Head Start, which serves families with children from birth to age 3. Psychological research on early intervention has proliferated, creating an expansive literature and sound knowledge base. Many research ideas designed and tested in the Head Start laboratory have been adapted in a variety of service delivery programs. These include family support services, home visiting, a credentialing process for early childhood workers, and education for parenthood. Head Start's efforts in preschool education spotlighted the value of school readiness and helped spur today's movement toward universal preschool.

The existing literature also provides some guidelines for program design. Arthur Reynolds suggests eight principles for designing effective interventions for children at risk for low academic achievement:

1. Target children and families who are at the highest risk of school difficulties.

2. Begin participation early and continue to the second or third grade.

3. Provide comprehensive child-development services.

4. Encourage active and multifaceted parent involvement.

5. Create a child-centered, structured curriculum approach.

6. Have small class sizes and teacher/child ratios.

7. Provide regular staff development and inservice training for certified teachers.

8. Evaluate and monitor the interventions systematically.

These principles suggest that although it may be useful to intervene before 3 years of age, interventions for preschool and for school-age children can also be effective. Thus, the first 3 years should not be emphasized at the expense of interventions aimed at older children. Second, the effects of early intervention have often been found to be larger for more disadvantaged children, which provides a rationale for targeting very high-risk children in particular. In addition to focusing on low-income children, it might be useful to target other aspects of disadvantage, such as lack of maternal education. Third, the most important aspect of quality is likely to be the nature of the interaction between the teacher and the child. Small group sizes, better teacher training, and other regulable aspects of quality all make such interactions more likely. Moreover, even rather loose regulation of these observable aspects of quality by Head Start has been shown to be effective in eliminating poor-quality programs.

Future of Risk Reduction Programs

What we know from the field of intervention and prevention programming is reflected in research on effective, risk-reducing, and resilience-building programs. Effective services provide contexts that both reduce the impact of risk factors and foster the development of new or existing protective factors. These programs build on inherent strengths within families, schools, and communities and enable these institutions to help children succeed. And more than just helping children, the best programs also support those who care for and provide services to these children, thereby enhancing their capacity to care. These programs address child development at a variety of stages, from prenatal care through postsecondary employment—stages that some would even say are too early or too late for appropriate intervention. These programs repeatedly demonstrate that resilience, rather than being solely dependent on individual characteristics, can be socially constructed.

Tiffany Berry and Elise Arruda

See also Head Start; Home Environment and Academic Intrinsic Motivation; Longitudinal Research; Social Class and Classism

Further Readings

Child Welfare League of America (The University of Illinois at Chicago), Reynolds, A. J., Wang, M. C., & Walberg, H. J. (Eds.). (2003). *Early childhood programs for a new century.* Washington, DC: CWLA Press.

Goldstein, S., & Brooks, R. B. (Eds.). (2005). *Handbook of resilience in children.* New York: Kluwer.

Groark, C. J., Mehaffie, K. E., McCall, R., & Greenberg, M. T. (Eds.). (2006). *Evidence-based practices and programs for*

early childhood care and education. Thousand Oaks, CA: Corwin.

Lerner, R. M., Wertlieb, D., & Jacobs, F. (Eds.). (2005). *Applied developmental science: An advanced textbook.* Thousand Oaks, CA: Sage.

Rosenthal Effect

The term *Rosenthal effect* is defined in its most general form as the effect of interpersonal expectations (i.e., the finding that what one person has come to expect from another can come to serve as a self-fulfilling prophecy). This concept is relevant to educational psychology in two distinct domains: the domain of research methodology (the experimenter expectancy effect) and the domain of learning and behavior (the teacher expectancy effect).

The concept of interpersonal expectation effects has been investigated in a wide array of settings, including the relationship between judges' expectations and their nonverbal behavior as they address the jury, and juries' subsequent verdicts of guilty or not guilty; the effects of managers' expectations for the performance of their employees, and employees' actual subsequent performance; and the effects of expectations of health care providers for their patients' subsequent health outcomes and patients' actual subsequent outcomes. Although interpersonal expectations have been studied in many domains, this entry gives primary attention to those domains of greatest relevance to students of educational psychology: the experimenter expectancy effect and the teacher expectancy effect.

Experimenter Expectancy Effect

The experimenter expectancy effect is one of the sources of artifact or error in scientific inquiry. Specifically, it refers to the unintended effect of experimenters' hypotheses or expectations on the results of their research.

Some expectation of how the research will turn out is virtually a constant in science. Social scientists, like other scientists generally, conduct research specifically to examine hypotheses or expectations about the nature of things. In the social and behavioral sciences, the hypothesis held by the investigators can lead them unintentionally to alter their behavior toward the research participants in such a way as to increase the likelihood that participants will respond so as to confirm the investigator's hypothesis or expectations. We are speaking, then, of the investigator's hypothesis as a self-fulfilling prophecy. One prophesies an event, and the expectation of the event then changes the behavior of the prophet in such a way as to make the prophesied event more likely. The history of science documents the occurrence of this phenomenon with the case of clever animals that were cued unintentionally to give correct answers in foot taps and in barking to questioners who believed the animals could respond with correct answers to, say, arithmetic problems.

The first experiments designed specifically to investigate the effects of experimenters' expectations on the results of their research employed human research participants. Graduate students and advanced undergraduates in the field of psychology were employed to collect data from introductory psychology students. The experimenters showed a series of photographs of faces to research participants and asked participants to rate the degree of success or failure reflected in the photographs. Half the experimenters, chosen at random, were led to expect that their research participants would rate the photos as being of more successful people. The remaining half of the experimenters were given the opposite expectation—that their research participants would rate the photos as being of less successful people. Despite the fact that all experimenters were instructed to conduct a perfectly standard experiment, reading only the same printed instructions to all their participants, those experimenters who had been led to expect ratings of faces as being of more successful people obtained such ratings from their randomly assigned participants. Those experimenters who had been led to expect results in the opposite direction tended to obtain results in the opposite direction.

These results were replicated dozens of times employing other human research participants. They were also replicated employing animal research subjects. In the first of these experiments, experimenters were employed who were told that their laboratory was collaborating with another laboratory that had been developing genetic strains of maze-bright and maze-dull rats. The task was explained as simply observing and recording the maze-learning performance of the maze-bright and maze-dull rats. Half the experimenters were told that they had been assigned rats that were maze-bright, and the remaining experimenters

were told that they had been assigned rats that were maze-dull. None of the rats had really been bred for maze-brightness or maze-dullness, and experimenters were told purely at random what type of rats they had been assigned. Despite the fact that the only differences between the allegedly bright and dull rats were in the minds of the experimenters, those who believed their rats were brighter obtained brighter performance from their rats than did the experimenters who believed their rats were duller. Essentially, the same results were obtained in a replication of this experiment employing Skinner boxes instead of mazes.

Teacher Expectancy Effect

If rats became brighter when their experimenters expected them to, then perhaps children could become brighter when expected to by their teacher. Accordingly, all of the children in the first study of the teacher expectancy effect were administered an obscure test of intelligence that was disguised as a test that would predict intellectual "blooming." The test was labeled "The Harvard Test of Inflected Acquisition." There were 18 classrooms in the school, three at each of the six grade levels. Within each grade level, the three classrooms comprised children with above-average ability, average ability, and below-average ability, respectively. Within each of the 18 classrooms, approximately 20% of the children were chosen at random to form the experimental group. Each teacher was given the names of the children from his or her class who were in the experimental condition. The teacher was told that these children had scored on the "Test of Inflected Acquisition" such that they would show surprising gains in intellectual competence during the next 8 months of school. The only difference between the experimental group and the control group children, then, was in the mind of the teacher.

At the end of the school year, 8 months later, all of the children were retested with the same test of intelligence. Considering the school as a whole, the children from whom the teachers has been led to expect greater intellectual gain showed a significantly greater gain than did the children of the control group.

Robert Rosenthal

See also Halo Effect; Teaching Strategies

Further Readings

Blanck, P. D. (Ed.). (1994). *Interpersonal expectations: Theory, research, and applications.* New York: Cambridge University Press.

Pfungst, O. (1965). *Clever Hans* (C. L. Rahn, Trans.). New York: Holt, Rinehart & Winston. (Originally published in 1911)

Rosenthal, R. (1966). *Experimenter effects in behavioral research.* New York: Appleton-Century-Crofts.

Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom.* New York: Holt, Rinehart & Winston.

Rosnow, R. L., & Rosenthal, R. (1997). *People studying people: Artifacts and ethics in behavioral research.* New York: W. H. Freeman.

RUBRICS

A rubric is a measurement and instructional tool that communicates instructor expectations to learners and uses explicitly stated criteria to categorize levels of performance regarding various skill levels, behaviors, product quality, and/or process quality. Rubrics are useful when instructors assess work in which learners have constructed knowledge and the work may not be easily scored using an answer key. For example, rubrics can be used to evaluate a performance of a task, a process such as problem solving, a written paper on a topic, a portfolio, and so on. In addition, a rubric enables both the instructor and the learners to identify desirable aspects of the work and areas needing improvement. A rubric often is a one- to two-page document containing a table or grid outlining criteria in categories that helps the learner identify the instructor's expectations for a task or project and assists the instructor in evaluating the learner's work. A rubric contains a Likert-type scale that helps to quantify performance decisions.

Rubrics may be generic or task-specific. A generic rubric is used to evaluate a process such as collaboration and may be applied across disciplines or content areas. A task-specific rubric is used for a clearly defined task and therefore has a narrower application. In addition, rubrics may be classified as holistic or analytic. Holistic rubrics enable the instructor to judge the overall product or process as a whole without focusing on separate components. This provides a quick, evaluative picture but does not give specific or detailed feedback to the learner. In other words, holistic rubrics require instructors to evaluate the

quality of learning through one performance level rather than multiple performance levels. In contrast, analytic rubrics have several indicator categories that instructors rate separately in order to differentiate performance levels within and among categories. Therefore, learners receive specific and detailed feedback that may be useful to them for improving performance.

A scoring rubric is used by instructors exclusively for the purpose of assigning grades to learners' work. In contrast, an instructional rubric helps learners to understand the instructor's expectations for an assignment and aids learners in focusing their efforts on the assignment. The same rubric may be used for both scoring and instructional purposes if desired. Instructors have found that instructional rubrics enable them to provide learners with more informative feedback without spending inordinate amounts of time. Also, instructional rubrics may help instructors to provide fair and unbiased evaluation of learners' work. In addition, multiple instructors can use the same rubric to attempt to have consistency in scoring. These rubrics can be used for learner self-evaluation and peer evaluation. Two key things to note when using instructional rubrics are that they should be provided to learners at the same time as the assignment and that the instructor should spend time explaining the rubric to the learners. Some instructors co-construct rubrics with the learners by first helping them to brainstorm criteria for which their work may be evaluated.

Rubrics can be obtained from a variety of print and Internet sources. If an appropriate rubric cannot be located, then an instructor may need to adapt an existing rubric or create a new one. The first step in creating a rubric is to examine examples of assignments to determine exemplary characteristics. These characteristics then can be grouped into categories. Next, levels of performance or gradations of quality are selected. Usually three to four levels are appropriate, although a rubric may contain more levels. One disadvantage of having three levels is that when learners use the rubric for self-evaluation, they may have a tendency to place their performance in the middle level. Levels may be labeled in a variety of ways, such as below average, average, and above average, or novice, apprentice, proficient, and distinguished. Point values also may be attached to each level of performance to help with scoring. Finally, indicators are identified for each category. Indicators

are brief statements that describe the particular characteristics of products that demonstrate that learning has occurred. It is important to use language in positive terms that learners, their family members, and other professionals will understand. Describing the best and worst levels of performance and then filling in the middle levels may be the easiest approach. Rubrics should be evaluated and refined continually if they are to be useful for all intended audiences.

In order for a rubric to be of high quality, it must have clear criteria as designated by the performance indicators. In other words, what is to be measured should be clearly defined. It is important to have enough criteria (e.g., three to seven categories), but not so many that the rubric is difficult to manage. Using rich, descriptive language enables learners to understand the indicators, and the indicators clearly should differentiate among levels of performance in measurable terms. Also, a rubric should emphasize positive attainment of desired performance as opposed to lack of attainment. A high-quality rubric should be valid (i.e., measure key features related to quality of performance) and reliable (i.e., consistent results for different users). Also, it should be sequenced to correspond with the steps in the performance. A high-quality rubric should have clear instructions for scoring. For example, it should be noted whether all categories have equal weight and scores from all categories should be added or averaged, or whether one or more categories should receive more weight.

There are many reasons why rubrics should be used in instructional settings. First, they help to make instructor expectations for learners' work clear and concrete. Second, rubrics provide learners with feedback regarding their strengths and weaknesses. Third, rubrics help learners to evaluate their own work and foster metacognition. Learners can use rubrics to monitor their own progress as they work on an assignment and use the rubric as a final checkpoint before turning in the assignment. The benefits of rubrics to instructors include establishing standards, linking assessment and instruction, providing a consistent and unbiased way of scoring work, and explaining those scores to learners and their family members.

Martha J. Larkin

See also Bloom's Taxonomy of Educational Objectives; Evaluation; Grading; Testing

Further Readings

Andrade, H. G. (2000). Using rubrics to promote thinking and learning. *Educational Leadership*, *57*(5), 13–18.

Andrade, H. G. (2005). Teaching with rubrics: The good, the bad, and the ugly. *College Teaching*, *53*(1), 27–30.

Bargainnier, S. (2004). Fundamentals of rubrics. In D. Apple (Ed.), *Faculty guidebook.* Lisle, IL: Pacific Crest.

Jackson, C. W., & Larkin, M. J. (2002). RUBRIC: Teaching students to use grading rubrics. *Teaching Exceptional Children*, *35*(1), 40–45.

Whittaker, C. R., Salend, S. J., & Duhaney, D. (2001). Creating instructional rubrics for inclusive classrooms. *Teaching Exceptional Children*, *34*(2), 8–13.

S

First you take a drink, then the drink takes a drink, then the drink takes you.

—Francis Scott Key Fitzgerald

Scaffolding

Scaffolding is a process in which support is provided to an individual so that he or she can complete a task that could not be completed independently. The support gradually is removed when the individual begins to demonstrate understanding of the task. The concept of *scaffolding* stems from the work of American psychologist Jerome Bruner and colleagues based on Lev Vygotsky's zone of proximal development. Vygotsky described the zone of proximal development as the distance between the actual developmental level where independent problem solving occurs and the potential developmental level where problem solving can occur with the guidance of an adult or more knowledgeable peer. Key to the zone of proximal development is social interaction and collaborative problem solving. Thus, the zone of proximal development bridges the gap between what an individual can learn and do independently and what he or she can learn and do with support. The scaffolding process in education bears similarities with the traditional definition of scaffolding, which is a temporary framework that supports workers and materials until a building is constructed or repaired to stand on its own. When scaffolding is used in instruction, learners receive support as needed and then the support gradually is removed as they achieve independence in task mastery.

Scaffolding instruction includes several essential elements that do not necessarily need to be followed in order. First, the teacher considers curriculum goals and standards along with student needs to select appropriate tasks. Second, the teacher works with students to establish a shared goal. This involvement may result in students who are motivated and invested in learning. Third, the teacher actively diagnoses student needs and understandings to ensure that students are making progress. Fourth, the teacher provides tailored assistance as needed through prompting, questioning, modeling, telling, or discussing. Fifth, the teacher helps students to remain focused on the intended goal by asking questions and providing clarification as well as offering praise and encouragement. Sixth, the teacher provides feedback in the form of a current progress summary and mention of specific behaviors that contributed to student success. Seventh, the teacher controls frustration and risk by creating an environment in which students feel comfortable taking risks with their learning without fear of penalty. Finally, the teacher gives students opportunities to practice the task in a variety of contexts and helps them to be less dependent on the teacher in order for them to internalize the task and eventually be able to perform it independently.

The following guidelines can help make the scaffolding process effective. The teacher can plan instruction by having students begin with tasks that they can perform successfully with little or no assistance in order for them to be aware of their strengths and feel

good about their abilities. Helping students to achieve success quickly may alleviate frustration. Then, more challenging tasks can be attempted with assistance. Peer acceptance is important to students, so it is important for the teacher to help students to appear like their peers when possible. Although practicing new and previously learned skills is essential, the teacher should recognize when too much practice may be contributing more to student frustration than to learning. The teacher should help the student with his or her current difficulties, redirecting the student's intentions only if he or she is not using an effective strategy for task completion. The teacher should watch for student clues as to when and how much assistance is needed. The assistance should be provided immediately to help the student perform the task, but also should be removed gradually as the student demonstrates task mastery.

The following lesson framework may help teachers provide scaffolding for their students. When presenting a new or difficult concept to students, it is likely that they will need more assistance at first. Therefore, the teacher may model how to perform the task while the students observe. Next, the teacher and students work together on the task. While the teacher demonstrates the task on the board, students may be performing the task on a handout while seated at their desks. The teacher may ask the students questions or prompt them to contribute to the class discussion on the task. At this point, the students gradually are taking some responsibility for their learning. When ready, students will work with a partner or a small cooperative group to perform the task. The teacher monitors the group work while gradually turning over more responsibility for the task performance to the students. Finally, the teacher has the students perform the task individually. The four stages of (a) teacher modeling, (b) teacher and students working together, (c) students working with a partner or small group, and (d) students working independently do not all have to occur in one class period. It could take 1 day to several weeks or more depending on how long a student needs to master the task.

Scaffolding is a challenging but beneficial process in instruction. One of the biggest challenges of effective scaffolding is that it is time consuming, particularly for one teacher in a classroom. Another is that judging the zone of proximal development for each student may be difficult (i.e., finding the area where a student needs help, but making sure that area is not

beyond the student's abilities). A third challenge is knowing student needs, interests, and abilities in order for the teacher to provide appropriate modeling. A fourth challenge may be making sure that the teacher gradually begins to fade or withdraw assistance as the student begins to demonstrate task mastery. Despite the challenges, scaffolding can be beneficial to students. It may provide a better chance of a student mastering the intended task. Second, the structured nature of scaffolding may ensure more time on task and efficiency in task performance. Third, scaffolding engages the student, motivates him or her to learn, and reduces student frustration. Fourth, it provides individualized and possibly differentiated instruction.

Martha J. Larkin

See also Constructivism; Effective Teaching, Characteristics of; Teaching Strategies; Zone of Proximal Development

Further Readings

Hogan, K., & Pressley, M. (Eds.). (1997). *Scaffolding student learning: Instructional approaches and issues.* Cambridge, MA: Brookline Books.

Larkin, M. J. (2001). Providing support for student independence through scaffolded instruction. *Teaching Exceptional Children, 34*(1), 30–34.

Larkin, M. J. (2002). *Using scaffolded instruction to optimize learning* (ERIC Digest No. 639). Arlington, VA: ERIC Clearinghouse on Disabilities and Gifted Education. (ERIC Document Reproduction Service No. EDO EC 02 17) Retrieved from http://eric.ed.gov/ERICDocs/data/ericdocs2/content_storage_01/0000000b/80/2a/38/e0.pdf

Lipscomb, L., Swanson, J., & West, A. (2004). Scaffolding. In M. Orey (Ed.), *Emerging perspectives on learning, teaching, and technology.* Retrieved from http://www.coe.uga.edu/epltt/scaffolding.htm

Verenikina, I. (1998). *Understanding scaffolding and the ZPD in educational research.* Retrieved from http://www.aare.edu.au/03pap/ver03682.pdf

Schemas

Although no universally agreed-upon definition of schemas exists, *schemas* are generally considered to be well-learned cognitive patterns of domain-specific information that are used as templates by individuals to help them explain, interpret, perceive, encode, and respond to complex tasks and experiences. Schemas also allow for predictions about what to expect in

future situations relevant to the particular schema. They create meaning from situations, data, and events by organizing and determining the patterns in complex sets of information. Schemas actually have a reciprocal relationship with data in that schemas may modify the meaning of information, but information or data may also lead to modifications in schemas. Both educators and counselors have interest in schemas because schemas help them understand how both informational and emotional learning occur.

Various types of schemas have been postulated, such as schemas about other people (including *role* and *person schemas*), one's self (*self-schemas*), the sequence of various events (*script schemas*), context (*place* or *location schemas*), and the meaning of data (*information schemas*). All of the various types of schemas facilitate the efficient understanding and interpretation of information by organizing and assigning meaning to that information. As a concrete example, suppose you heard someone talking about seemingly highly disparate pieces of information, such as arranging things into groupings; making decisions about color; dealing with tedium; deciding about capacity and facilities to employ; avoiding mistakes; timing of mechanisms; sorting types; determining what could not be dealt with by one's current equipment, which necessitates outsourcing; setting temperatures in such a way as to avoid catastrophe; dealing with voluminous output; enlisting aid from others; and making measurements of necessary additives. This list might sound rather convoluted, meaningless, and difficult to remember unless you were first told that the pieces of information all concerned "doing laundry."

During the initial learning process, deliberate construction of schemas requires the use of significant amounts of working memory (WM) resources. Working memory represents the brain's capacity to temporarily hold limited amounts of information while manipulating that information. However, with practice and repetition, the use of schemas constructed during the learning process becomes virtually automatic. Thus, the development of schemas allows for a substantial reduction in required WM resources as the schemas direct and guide individuals' attention and focus. The result is often an increase in expertise or skill level within a particular knowledge domain.

Schemas also provide an overall executive guidance system during high-level cognitive processing. Without this guidance (or without external instruction), individuals often default to weak problem-solving strategies, such as trial and error and means-ends analysis. Strategies such as these can be both time consuming and inefficient, and thus interfere with the construction of new schemas because of the workload imposed on WM resources.

Schemas are stored in long-term memory (LTM), which is virtually unlimited in both its capacity and duration and allows individuals to process, organize, and retrieve vast reservoirs of knowledge. Once schemas are formed and stored in LTM, working memory is freed up to process, interpret, and ultimately store new schemas into LTM. By and large, when schemas are needed in WM, they are dealt with as a single piece of information, although they contain a rich array of data. Therefore, many educators consider schema formation to be an important focus for instructional design.

The concept of schema acquisition has significant applicability to instructional design. Often, it is recommended that one assess students' current knowledge and activate relevant schemas prior to teaching new information. Analogies, metaphors, and comparisons can be useful in this regard. Schema acquisition is also facilitated by providing already completed, worked-out examples prior to asking students to work out entirely new problems all on their own. Consideration should also be given to the cultures from which students have come, in that culture provides a backdrop and context for their existing schemas. Thus, asking a student to learn new material in a highly individualized, competitive atmosphere might impede schema development for a student who comes from a culture that emphasizes cooperation, mutuality, and group membership.

Optimally, one wishes to develop schemas that have maximal flexibility and applicability to a wide range of situations and contexts. Therefore, it is often recommended that one attempt to facilitate the development of schemas that are not overly constricted to specific situations. Such flexible schemas may be acquired through exposure to similar knowledge that is applied to a range of contexts other than the one in which learning originally occurred.

The development of more flexible schemas also has relevance to the emotional responses of students. Thus, students' inflexible *self-schemas* (which have also been referred to as *early maladaptive schemas*) tend to be associated with stronger negative affect, which can readily interfere with learning. By using a schema-based approach to students' dysfunctional schemas and subsequent behaviors, the school counselor can create an atmosphere based on a nonpathological model and

focus on the ways in which a student has learned inflexible, intensely emotional responses to certain types of situations. Inflexible schemas can be examined and moderate, flexible schemas developed to take their place. Such moderate, flexible schemas are likely to be associated with less intense emotions and improved learning.

Charles H. Elliott and Tam Chandler

See also Cognitive Behavior Modification; Cognitive View of Learning; Culture; Long-Term Memory; Working Memory

Further Readings

Chen, Z. (2006). Generalization and transfer of problem solving strategies. In A. V. Mittel (Ed.), *Focus on educational psychology* (pp. 217–234). New York: Nova Science.

Elliott, C. H., & Lassen, M. K. (1997). A schema polarity model for case conceptualization, intervention, and research. *Clinical Psychology: Science and Practice*, *4*(1), 12–28.

Kalyuga, S. (2006). Instructing and testing expertise: A cognitive load perspective. In A. V. Mittel (Ed.), *Focus on educational psychology* (pp. 53–104). New York: Nova Science.

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist*, *38*(1), 23–31.

SCHOOL COUNSELING

School counseling, a crucial component to students' achievement, is a comprehensive program that facilitates students' academic, career, and personal/social development within the school setting. Professional school counselors have a minimum of a master's degree in school counseling. In order to facilitate the development of all students, professional school counselors implement a wide range of therapeutic interventions, including classroom guidance lessons on topics such as anxiety management and bully prevention, group and individual counseling, career testing and planning, parent and teacher consultation, and advocacy for systems change. Research has shown that these school counseling services improve students' academic success. School counseling is an important topic in educational psychology because it promotes students' academic, career, and personal/social achievement in the educational settings of elementary, middle, and high schools. This entry provides details of school counseling by (a) reviewing the history of school counseling; (b) explaining the American School Counseling Association's National Model for school counseling; and (c) describing key components of a comprehensive school counseling program including guidance curriculum, counseling, career development, consulting, coordination of resources, leadership and advocacy, promotion of a safe and respectful climate, accountability, management of legal and ethical issues, and professional development.

History

The forerunner of modern school counseling was vocational guidance, a preventive educational approach that taught students how to deal with life events. In 1908, Frank Parsons, known as the "Father of Guidance," founded Boston's Vocational Bureau, where he helped young people with career decisions. Concurrently, in 1907, Jesse B. Davis implemented weekly vocational and moral guidance lessons during English classes in Grand Rapids, Michigan, which led to a systematized guidance program in the public schools. The school guidance movement strengthened as Harvard University began education courses for counselors in 1911; the National Vocational Guidance Association was established in 1913; and the Smith-Hughes Act of 1917 funded vocational education in public schools.

Over the 100-year history of school counseling, program focus and duties evolved in response to changing trends and needs. In the early 1900s, school counselors focused on scheduling student courses that would lead to careers needed in the Industrial Revolution. In the 1910s, psychometrics became another focus when guidance workers used the military's Army Alpha and Army Beta intelligence tests to identify highly capable students. In the 1920s, secondary school guidance personnel were trained similar to college personnel because of limited training programs and thus acquired some administrative and disciplinary duties, similar to college deans of students. In the 1930s, school guidance personnel followed E. G. Williamson's approach of enhancing normal adjustment by helping individuals to set goals and teaching them needed skills. In the 1940s, Carl Rogers's nondirective emphasis of listening and accepting clients without judgment resulted in school

counselors providing client-centered counseling to students, rather than just guidance.

In the 1950s, after the Soviet Union launched its first space satellite, *Sputnik I*, the United States funded the National Defense Education Act (NDEA). As a result, school counselors focused on student career testing to channel students with high math and science abilities into college. In addition, NDEA funded elementary school counseling so that talented elementary students could be identified. The 1960s group encounter movement influenced school counselors to offer small group counseling. Concurrently, C. Gilbert Wrenn advocated that school counselors expand their focus to the developmental needs of all students rather than just the top or bottom percentage. Hence, the focus shifted to the developmental guidance approach of promoting positive individual growth and preventing problems.

Despite declining school enrollment and economic problems in the 1970s and 1980s, school counselors continued to expand the developmental guidance focus to students' self-understanding and adjustment as well as career development from kindergarten through 12th grade. In 1998, Norman Gysbers and Patricia Henderson published *Developing and Managing Your School Guidance Program*, which provided guidelines for a comprehensive developmental guidance program. In the 1990s, multiculturalism became prominent in counseling, inspiring school counselors to pay more attention to the varying needs of students from different ethnic and socioeconomic groups.

Spanning from the end of the 1990s through the new millennium, concerns of school violence, bullying, and crises emerged due to a rash of school shootings and the September 11, 2001, terrorist attacks. School counselors responded by focusing on bully prevention and developing crisis counseling teams. Because of limited resources, the prominence of managed care within the health care system, and accountability requirements of the 2001 No Child Left Behind policy, school counselors began to focus on accountability by providing data to prove that interventions led to student success.

American School Counseling Association National Model

In order to stabilize the changing focus of school counseling, in 2003, the American School Counseling Association (ASCA) developed a national model to provide consistent, comprehensive guidelines for school counseling programs and professional school counselors' duties that would promote success for all students throughout the country. A detailed description of the ASCA National Model is provided in *The ASCA National Model: A Framework for School Counseling Programs, Executive Summary*.

> The ASCA National Model supports the school's overall mission by promoting academic achievement, career planning and personal/social development. It serves as a framework to guide states, districts and individual schools in designing, developing, implementing and evaluating a comprehensive, developmental and systematic school counseling program. (ASCA, 2003, p. 1)

This approach to school counseling programs benefits students, parents, teachers, administrators, and the overall community. It is an integral part of each student's achievement. It is systematically delivered to every student and is not just for high achievers or at-risk students.

Systematic delivery of the ASCA National Model encompasses four interrelated components: foundation, delivery system, management systems, and accountability. The foundation is composed of (a) beliefs and philosophy on which all personnel agree; and (b) a mission statement that highlights the program's purpose, which aligns with the school and district's mission. The delivery system entails four methods needed to systematically deliver the school counseling program to all students. The first method of the delivery system is guidance curriculum. The curriculum consists of structured classroom lessons that provide knowledge and skills at the appropriate developmental level for kindergarten through 12th-grade students. The second method is individual student planning in which professional school counselors meet with individual students to help them identify goals and future plans. The third method is responsive services to meet individual students' immediate needs through counseling, consultation, referral, peer mediation, or provision of information. The fourth method of the delivery system is systems support via administration and management of the total counseling program.

The management system, the third component of the ASCA National Model, incorporates organizational processes to make sure the counseling program is aligned with the school's needs. Agreements about the school counseling program's organization and

goals are negotiated with school administrators. An advisory council of students, parents, teachers, counselors, administrators, and community members is established to review counseling program results and make recommendations. Data are used to decide what activities are needed to promote students' academic, personal/social, and career achievement. Action plans are developed to achieve every desired competency and result. These action plans describe in detail the competencies addressed, activity components, data indicating the need for the activity, time line, responsible party, evaluation methods, and expected outcome. School counselors' time should be carefully guarded so that 80% of their time is spent on direct service contact with students. Their duties should be limited to program delivery rather than noncounseling activities. Calendars are developed and published so that students, parents, teachers, and administrators know the school counselor's schedule of activities.

Accountability is the final component of the ASCA National Model. To hold professional school counselors accountable, data are used to link school counseling activities to student achievement. Results reports ensure that programs were implemented, analyzed for effectiveness, and modified for activity and program improvement. Immediate, intermediate, and long-range reports are shared with stakeholders. School counselor performance standards are used to evaluate the school counselor and school counseling program. Finally, program audits are conducted to guide future action within the program.

ASCA's national standards for student academic, career, and personal/social development outline competencies that students will obtain or demonstrate as a result of the school counseling program. Regarding academic development, students will (a) acquire attitudes, knowledge, and skills for effective learning; (b) complete school with academic essentials needed for postsecondary options; and (c) understand the relationship between academics and career as well as life in the community. Regarding career development, students will (a) acquire skills to investigate work options and self so that they can make informed career decisions; (b) utilize strategies to achieve future career goals with success and satisfaction; and (c) understand the relationship between personal characteristics, education, training, and career options. Regarding personal social development, students will (a) acquire knowledge, attitudes, and skills to understand and respect self and others; (b) learn to make decisions and establish

and achieve goals; and (c) develop safety and survival skills. Professional school counselors must ensure that their programs help students accomplish each of these competencies.

In order to implement school counseling programs as described above, school counselors' duties and responsibilities must be delineated clearly. Although professional school counselors are team players, they cannot be fully effective when they perform noncounseling administrative or clerical activities such as developing the master school schedule, acting as testing coordinators, providing coverage for classrooms and detention rooms, disciplining students, or performing clerical duties. Rather, appropriate school counseling responsibilities include, but are not limited to, the following:

- Implementing individual student academic programs
- Analyzing and conveying cognitive, aptitude, and achievement tests
- Counseling students with disruptive problems
- Collaborating with teachers on stimulating guidance curriculum lessons
- Analyzing grade point averages and achievement tests
- Interpreting student data
- Consulting with teachers on study hall management
- Collaborating with the school principal on resolving student issues

A detailed description of school counselors' essential functions follows.

Key Components

Developmental Classroom Guidance Curriculum

Professional school counselors implement a developmental classroom guidance curriculum to all students in an effort to prevent problems in students. The curriculum addresses common concerns that are identified by needs assessments of students, faculty, and parents as well as national standards related to academic, personal/social, and career development. Professional school counselors present structured, planned lessons to a large group of students to meet students' developmental needs. Topics of typical guidance lessons include the following:

- Academic skills support
- Organizational, study, and test-taking skills

- Education in understanding self and others
- Coping strategies and crisis management skills
- Peer relationships and effective social skills
- Communication, problem solving, decision making, conflict resolution, and bullying prevention
- Career awareness, exploration, and planning
- Substance abuse education
- Multicultural/diversity awareness

Although the topics may remain the same across school levels, developmental approaches vary based on school level as toys and puppets may be used in elementary school whereas games and role-playing may be used in secondary schools.

Counseling

When students need more intensive help beyond classroom guidance lessons, school counselors provide small group counseling and individual counseling. During small group counseling, professional school counselors meet with two to eight students at a time to provide therapeutic intervention that meets participants' individual and common goals. In elementary schools, typical goals include improving social skills and general behavior, adjusting to family changes of divorce or death, or resolving underlying personal problems that interfere with academic success. In secondary schools, typical goals include student success skills and career decision making. Professional school counselors employ group counseling skills of facilitating open communication, linking group members together, providing feedback on interactions, confronting disruptive behavior, and encouraging positive interaction with other members.

When providing individual counseling to a student, professional school counselors may follow a particular counseling theory such as person-centered (Carl Rogers); Adlerian (Alfred Adler); reality therapy (William Glasser); or cognitive behavior therapy (Aaron Beck, Donald Meichenbaum, G. Terence Wilson). However, many professional school counselors are eclectic and use skills from various theories to meet an individual student's needs. In secondary schools, solution-focused counseling (Steve de Shazer, John Murphy) provides a quick and simple approach of helping students focus on small changes and reasonable goals. The steps are as follows:

1. Establish a cooperative, change-focused relationship.

2. Develop small, meaningful goals.

3. Encourage students to do more of what works and use their resources.

4. Change the "doing" by trying a different strategy or changing location, frequency, intensity, or sequence of behavior; and change the "viewing" by seeing it from a different perspective.

5. Evaluate and maintain progress.

In elementary schools, play therapy or "counseling with toys" (Garry Landreth) allows students to express their feelings, thoughts, and behaviors and resolve conflicts through their natural medium of communication, play. Professional school counselors create a safe environment with specially selected toys and provide therapeutic responses of tracking play behavior, reflecting feelings and content, returning responsibility, encouraging, building self-esteem, setting therapeutic limits, facilitating understanding, and expanding the meaning of the child's play. Numerous research studies have demonstrated that play therapy decreases children's behavior problems and increases their mental health, which in turn helps them succeed at school.

Career Development

Career development and guidance in the schools is based on the goal that each student will complete high school prepared for a variety of workplace and postsecondary options, including 2- and 4-year colleges, technical schools, or military service. Career development begins in preschool and continues beyond high school. According to Edwin Herr and Stanley Cramer, children often make a tentative commitment to a vocation in the first 6 years of school. Elementary school counselors focus on career awareness that transcends socioeconomic levels and gender roles. Common career awareness activities are career days, in which parents and community members present information on their careers; field trips to local industries, banks, hospitals, stores, and so on; and videos of a wide range of careers. Middle school counselors focus on career exploration. They help students explore work options in light of their own strengths, weaknesses, interests, talents, and skills. Middle school students are taught how to use computerized career information delivery systems to conduct assessments and occupational searches and obtain occupational and educational information. High school counselors focus on career planning by linking students' interests and skills with occupational and educational options. They

provide students with information, emotional support, reality testing, and planning strategies. High school students benefit from experiential activities such as attending a career fair, visiting a college campus, completing financial aid applications, and participating in a career internship. High school counselors must also give specific care and information to students who are potential high school dropouts.

Consultation

Consultation is the process in which professional school counselors (consultants) assist teachers, parents, administrators, and community members (consultees) with problems related to a student and the system (client system). Typical consultation issues include teachers' management of students' classroom behavior and teaching strategies for students with disabilities, parents' discipline of unruly children and motivating their children to do homework, administrators' concerns regarding low-performing teachers or decreasing violence and prejudice at school, and community members' desire to mentor students or provide resources.

Michael Dougherty described a generic model of consultation with four stages. During the entry stage, the school counselor (consultant) enters the relationship with the consultee by discussing roles, agreeing for action, and communicating confidentiality. In the diagnosis stage, the consultant assesses the consultee characteristics that may have an impact on the problem, the client characteristics, and the environment. Together, they clearly define the problem and develop measurable and realistic objectives. In the implementation stage, the school counselor and consultee select acceptable strategies and identify resources within the system. The strategies are implemented, the consultant observes and gives feedback, and strategies are adapted and adjusted as needed. In the disengagement stage, the consultant and consultee evaluate the effectiveness of the strategies in resolving the problems. Finally, they terminate the consultation by discussing the ending, gradually reducing involvement, and conducting a summary conference.

Coordination and Resource Management

Professional school counselors coordinate and manage resources to meet various needs of students, teachers, and parents, such as the need for a mentor, clothing, classroom supplies, computers, and so on. Clearly, professional school counselors cannot directly meet all students' needs alone. However, they can coordinate resources provided by parents, teachers, human service agencies, community members, and business partners. For example, the Parent Teacher Association may be able to organize a clothing closet, tutoring, and parent education classes. Business partners may be willing to provide career mentors and funding for new computers. Professional school counselors maintain resource lists and contacts with human service agencies such as after-school boys and girls clubs, a domestic violence shelter, public health clinics, and so on. Training peer mediators to help resolve student conflicts is also an effective use of resources. In doing so, discipline referrals are decreased and a positive attitude toward school is increased.

Leadership and Advocacy

Professional school counselors are collaborative leaders and advocates within the school system. As part of the leadership team, professional school counselors actively participate on school improvement teams to create a system in which all students can experience academic, career, and personal/social success. Professional school counselors are also leaders by working in partnership with principals and other key stakeholders, creating a positive school climate, conducting staff development for teachers, developing high aspirations in students, and using technology to track data.

Professional school counselors are advocates who question the status quo, challenge rules and regulations that deny student access, protest changes that hinder underrepresented groups, empower people who need strength, and promote needed changes in the system. They work to close the achievement gap for low socioeconomic students and minority students of color who have lacked support in achieving at the same academic level as majority White students. They advocate for these students by removing barriers that prevent their achievement in rigorous academic courses such as calculus, chemistry, or Advanced Placement English. Professional school counselors also advocate for safer school environments for students and teachers persecuted and oppressed because of their sexual orientation. For example, they may sponsor a Gay/Straight Alliance student group.

Promotion of a Safe and Respectful School Climate

Given numerous acts of school violence and federal mandates for safe and drug-free schools, professional school counselors must promote a safe and respectful school climate. Bullying is the most common form of school violence. Sexual harassment is another major concern of the majority of students. According to Carolyn Stone and Carol Dahir, risk factors for this type of violent behavior include alienation, depression and anxiety, destructive behavior, gang involvement, bias and prejudice, and use of drugs. In order to address this, professional school counselors provide guidance lessons on respect and bully prevention, train teachers to be sensitive to alienated and troubled students, ask all stakeholders to contribute to a respectful environment, and build positive relationships with students who are victims and perpetrators. In addition, the U.S. Department of Education stated that positive school climates can be created by (a) building a solid foundation for all children, (b) identifying and providing intensive interventions for at-risk students, (c) involving community members and agencies in creating a safe school environment, and (d) integrating character education across the content area. Professional school counselors are leaders in convening stakeholders in developing a schoolwide anti-bullying policy that includes clear definitions and disciplinary actions. This policy should be frequently communicated to students, parents, teachers, and administrators. Both teachers and parents should be trained in early warning signs and involved in school violence prevention programs.

Accountability

Rather than asking, "What do school counselors do?" the more important question is, "How are students different *because* of what school counselors do?" To answer this question, professional school counselors highlight the success of their school counseling program through result-based accountability. For example, professional school counselors may report that as a result of their anti-bullying program, 90% of the student population can recite the school anti-bullying policy, and as a result of their peer mediation program, 30% more students used peer mediators to resolve conflicts. Professional school counselors also use critical data elements, such as attendance rates, discipline referrals, graduation rates, and standardized test scores, to demonstrate the effectiveness of their programs. Stone and Dahir recommended that professional school counselors use the following six-step accountability process:

1. Connect the school counseling program with the school's mission and improvement plan.

2. Identify and examine critical elements of available data to the school's mission.

3. Analyze critical data elements to establish baseline and set goals.

4. Involve stakeholders in defining target results and strategies to improve the data.

5. Reanalyze the data to see if target goals were met and to reassess strategies.

6. Educate internal and external stakeholders by disseminating the data.

Management of Legal and Ethical Issues

Professional school counselors must manage numerous legal and ethical issues. Laws dictate minimum standards of behavior tolerated by society, whereas ethical standards, established by the ASCA, represent ideal aspirations of counselors. Frequently, laws and ethical codes contradict each other. For example, professional school counselors are required by state law to report child abuse to designated authorities. This law supersedes the ethical standard of maintaining confidentiality (i.e., respecting clients' right to privacy). In this example, professional school counselors balance adherence to laws and ethics by providing students with informed consent (i.e., ensuring that students understand the legal limits of confidentiality before counseling begins).

Other pertinent legal codes that professional school counselors must follow are (a) Public Law (P.L.) 94–142, the Education of All Handicapped Children Act of 1975, which ensures the free, appropriate public education of all children in the least restrictive environment; (b) Title IX of the Education Amendment Acts of 1972, which prohibits discrimination on the basis of gender; (c) Title II of the Education Amendment Acts of 1976, which ensures equal access to vocational education to both men and women; (d) P.L. 93–380, the Family Educational Rights and Privacy Act of 1974 (Buckley Amendment), which gives parents and students access to their entire

educational record; and (e) the No Child Left Behind Act of 2001, which states that all students will reach high standards of proficiency or better in reading/language arts and mathematics and will be taught by highly qualified teachers. Professional school counselors must also be familiar with state laws regarding issues such as privileged communication, HIV reporting, parental consent for abortions, and so on.

The ASCA has established ethical standards for professional school counselors (available at http://www.schoolcounselor.org) that address issues such as confidentiality, dual relationships, student records, parent rights, professional relationships, and so on. Common ethical dilemmas that professional school counselors encounter are (a) maintaining students' confidentiality while respecting parents' rights for information, (b) stressing the importance of confidentiality to students in group counseling but informing them it cannot be guaranteed, (c) following the duty to warn or protect when others report that a student made suicidal or homicidal ideations, (d) responding to a student's request to get an abortion when she states that her parents are against it, and (e) intervening when a student alleges sexual harassment or racial discrimination by a teacher who states that it did not happen.

In order to resolve such ethical dilemmas, professional school counselors use an ethical decision-making model that typically includes the following steps:

1. Identify the problem by gathering information and examining facts.

2. Consider state law, the ASCA Code of Ethics, standard of practice, school district policy, and the best interests of the client.

3. Determine the nature and dimensions of the problem in light of moral principles of autonomy, beneficence (doing good), nonmaleficence (doing no harm), justice, and loyalty.

4. Generate a potential course of action.

5. Consider the consequences of all options.

6. Evaluate the selected course of action and seek consultation.

7. Implement the course of action.

Professional Development

Because professional school counselors are entrusted with a demanding role and numerous responsibilities, they must renew themselves through professional development and personal wellness. They regularly participate in professional development activities such as joining professional organizations (e.g., the American School Counseling Association and the American Counseling Association), reading professional journals (e.g., *Professional School Counseling* or *Journal of Counseling and Development*), attending state and national conferences, obtaining advanced training via workshops or graduate courses, and seeking supervision from a seasoned school counselor. Professional school counselors also promote their personal wellness by nurturing family relationships, taking vacations to rest and relax, developing friendships with positive colleagues, exercising regularly and eating healthy, engaging in spiritual rituals such as prayer or meditation, and reading for fun.

Professional school counselors fulfill a crucial role in promoting students' academic, personal/social, and career achievement. They possess positive attitudes, knowledge, and skills to provide developmental guidance lessons, counseling, consultation, coordination, leadership, advocacy, and accountability. School counseling is an exciting, demanding, and fulfilling profession for energetic, compassionate, and organized professionals who desire to have a positive impact on students, school personnel, and the community.

Jennifer Baggerly

See also Eating Disorders; Gifted and Talented Students; Individual Differences; Personality Tests

Further Readings

American School Counseling Association. (2003). *American School Counseling Association National Model: A framework for school counseling programs.* Alexandria, VA: Author. Retrieved September 12, 2007 from http://www.schoolcounselor.org

Baker, S. B., & Gerler, E. R. (2004). *School counseling for the twenty-first century* (4th ed.). Upper Saddle River, NJ: Pearson Education.

Gladding, S. T. (2004). *Counseling: A comprehensive profession* (5th ed.). Upper Saddle River, NJ: Pearson Education.

Herr, E. L., Cramer, S. H., & Niles, S. G. (2003). *Career guidance and counseling through the lifespan* (6th ed.). Boston: Allyn & Bacon.

Stone, C. B., & Dahir, C. A. (2006). *The transformed school counselor.* Boston: Lahaska.

Stone, C. B., & Dahir, C. A. (2007). *School counselor accountability: A measure of student success.* Upper Saddle River, NJ: Pearson Education.

SCHOOL DESIGN

School facilities—the physical settings that contain and support teaching and learning—play an important and underappreciated role in providing effective conditions for learning. National and international reports on the current state of the physical infrastructure of schools make it clear that the deteriorating quality of the physical environment is affecting the quality of educational delivery. Because of their age, now well over an average of 42 years, school buildings do not always contain what is now considered the essential components for a good learning environment. There is arguably an ample body of evidence that school environments influence a number of student behaviors and attitudes that influence educational outcomes. Although there is some skepticism about the relationship between building condition and educational outcomes, recent studies have found significant correlations between building condition and academic achievement. Systemic, schoolwide educational reforms place urgent demands on school buildings originally designed for an industrial-age disciplinary mass institution. Other factors include the needs of a growing ethnic diversity of the student body, ever-present overcrowding, community education, and information and communications technology.

The 1960s witnessed one of the most dramatic educational reform movements in U.S. history with the experimentation of open education, community education, middle schools, and alternative and magnet schools, with innovations in school design led by the Educational Facilities Laboratory, or EFL. One of EFL's most influential innovations was the development of the open plan school design, a concept that influenced the design of thousands of schools from the late 1950s through the early 1970s. Schools were planned with large, open, flexible spaces adaptable to team teaching and small group and individualized instruction that characterized open education. Open education, it was argued, provided more educational opportunities for children, provided freedom and autonomy for self-directed study, required less guidance by the teacher, and helped foster self-responsibility.

Almost immediately, however, teachers complained of noise and visual distractions in these open plan schools. Hundreds of educational research studies were performed to determine the validity of open plan schools with inconclusive and controversial results. This period marks the most extensive empirical research conducted on educational environments and as such is a starting point for understanding the role of physical environments, good or bad, on teaching and learning.

The middle school concept, first conceived of in the 1960s, was a philosophy that challenged the junior high school model and advocated for the developmental needs of young adolescents—a balance of child-centered, supportive instruction of the elementary school with the subject-oriented teacher specialization of the high school. Middle school teachers formed a small interdisciplinary team that comprised a family of between 100 and 120 students. Gaining popularity in the 1980s and 1990s, the middle school concept spawned a whole new generation of school design that attempted to group students in "families" contained in "pods" or "houses." Pod plans were first developed in the 1960s, whereas the house plan has a more recent history, being most fully developed in the late 1980s. House plans, it is argued, foster a sense of community for academics while providing larger common spaces such as libraries; media centers; administrative functions; and gymnasia and special programs such as art, music, computer instruction, and language arts. The house may include anywhere from four to eight self-contained classrooms oriented toward a centralized resource center and supported with a specialized classroom, teacher offices, small seminar rooms, and other support spaces. Currently, the house plan concept is being applied to secondary environments as well, as a response to advances in self-directed learning, interdisciplinary instruction, and the desire to form smaller learning communities in very large high schools. The residential metaphor of the house has been extended to include the "neighborhoods" and "main streets" in high school design, thereby extending the notion of a community of learners.

School designs continue to respond to changes in education. Some innovations in school design, such as self-directed learning environments, are just now finding acceptance in education. Other innovations, such as wireless computing, are completely unprecedented and, in many ways, are driving a new set of changes. The characteristics of a new 21st-century school facility are

starting to emerge as a school that intentionally supports academic achievement, is personalized to the needs of the learner, is designed for multiple intelligences of learners, authentically engages the community, is flexible to adapt to multiple pedagogies, is sustainable and high performance, acts as a three-dimensional textbook, supports outdoor learning, is technology-rich on site and at a distance, and is continuously monitored through evaluation. Several of the emerging and interconnected themes in school design that are covered here include the advent of school designs for smaller learning communities, collaborative learner-centered environments, the role of community partnerships and joint-use agreements in school planning, the impact of technology on school design, and a trend toward high-performance schools.

There is an intense interest and rapid growth in the creation of smaller school environments in the United States, either by building smaller school buildings or by restructuring larger school buildings into a number of schools-within-a-school.

The literature on the effects of school size on a variety of school outcomes indicates that participation in school activities, extracurricular activities, student satisfaction, number of classes taken, and community employment have all been found to be greater in small schools relative to large schools. Smaller schools may have a positive influence on achievement, school climate, and student connectedness, thereby reducing disciplinary problems, incidents of vandalism, truancy, drug use, and drop out rates. Defining what constitutes the optimal organizational size and structure for schools-within-a-school is still an open question. School size research most often refers to the size of the student body, with large being defined as anywhere from 1,000 to 2,000 or more students in secondary schools, and small being considered anywhere from 100 to 600 students in elementary and even secondary settings.

The school design that emerged out of the middle school movement, the house plan (e.g., putting a small family of 100–120 students and their teachers into groupings of classrooms), serves as the core model for what is now called the "neighborhood" plan for the learning "community." The goal is the same, to break down the scale of the school organization into manageable groups of learners to create a natural sense of belonging, connectedness, and caring. An important characteristic of schools-within-a-school is that each school has a distinct administrative entity. Several models that are emerging include vertical houses, ninth-grade houses, and special curriculum houses. The vertical house plan assigns several hundred students and their teachers, grades 9–12, to a single house. The ninth-grade house plan provides an environment for ninth graders to ease the transition into high school. The special curriculum house plan organizes students into houses based on special interests or needs.

Another educational trend to which school designers are beginning to respond is classroom environments that emphasize active, self-directed, project-based, collaborative, or cooperative learning strategies over traditional, lecture-oriented, discipline-focused, teacher-centered instruction. Collaborative learning is a pedagogy that prepares learners for the changing learning expectations in the real world through an active learning process that teaches thinking critically, solving problems, working in teams, negotiating, reaching consensus, using technology, and taking responsibility for one's own learning. Self-directed learning, or the process of learning on one's own, has long been seen as a natural and primary mode of adult learning in which personal growth is the primary goal. Growth outlines the stages of the student–teacher relationship leading to self-directed learning: from dependence on the authority of the teacher; to interest, with the teacher as motivator; to involvement, with the teacher as facilitator; to fully self-directed, with the teacher as consultant or delegator. Public school educators who espouse self-directed learning tend to use teaching tools such as individualized, or personalized, learning plans; contracts; and advisories to frame either individual learning method.

School planners and architects have developed a variety of school designs that support personalized, self-directed learning. Wolff identifies a number of design features that may support learning in groups, including variable-sized spaces that are easy to change to support several learning activities within the same space and to encourage integration of courses and programs; individual work spaces that can be personalized, thus providing a sense of ownership and responsibility for one's own learning; and faculty team spaces with adjacent material preparation areas and meeting space that encourages team teaching, mentoring of faculty, and collaboration. Functional spaces could include presentation spaces for individuals and teams to demonstrate their learning and share knowledge acquired within the larger learning community; classroom spaces for direct instruction of concepts, content, and skills; process galleries or studios that allow for the display of ongoing projects to showcase concept

development; project space that provides a variety of work surfaces, storage space, and access to technology to encourage critical thinking, problem solving, and teamwork; a home base for the gathering of learners and faculty to seek assistance and resources or hold group discussions; informal, nonclassroom learning spaces such as study spaces, lounges, and outdoor spaces to provide areas for socializing and serendipitous meetings that can foster creative thought and solutions to problems; and a collaboration incubator or idea generation space to support creativity, teamwork, and prototyping of concepts, activities that can also encourage the involvement of local employers in the development of projects.

Another trend, that of high-performance schools (e.g., green, or sustainable), aims to support students and teachers to perform at their highest potential by being comfortable, healthy, safe, environmentally sound, and economically operational. The concept of the high-performance school integrates the environmental sustainability and building performance movements, which are concerned ultimately with the sustainability of the human species and the impact of our species on the planet. Ninety percent of our lives are spent indoors, with students and teachers spending more than 2,000 hours a year in school buildings alone.

Designing high-performance buildings can have a direct influence on learning through designing and maintaining a healthy and safe indoor environment, providing acoustical comfort, optimizing natural light, sharing community resources, and using the school building as a learning tool.

Children are the most vulnerable population with regard to the quality of the indoor environment because of their physical and cognitive developmental needs. Problems with indoor air quality and ventilation in concert with temperature and humidity can have profound effects on student and teacher physiology and psychology. Poor indoor air quality has been associated with the rise of a variety of pollutants, mold growth, and chemicals that can spread disease and result in a worst-case scenario of sick building syndrome. Increases in carbon dioxide simply from human respiration as well as other contaminants, if not ventilated, can cause sleepiness, increased mental fatigue, and decreased alertness, leading to a lack of concentration and focus on learning. According to the American Lung Association, more than 4.8 million school-aged children suffer from asthma in the United States. Prolonged exposure to these contaminants in a poorly ventilated building may lead to allergies, headaches, dizziness, and asthma, resulting in increased absenteeism. It is estimated that students are absent nearly 10 million school days/year due to illness caused by indoor environmental quality problems in schools. Providing proper ventilation, by designing increased ventilation rates as well as providing operable windows in every space, alone may improve many of the health concerns as well as performance and teacher retention concerns in many schools. Acoustics are well known to influence student performance. Children who are still developing their speech perception have a more difficult time hearing than do adults, with nearly 20% of children experiencing hearing loss. Heating, ventilation, and air conditioning equipment is known to be one of the biggest contributors of noise pollution in buildings. Excessive noise both inside (other students, other classrooms, hallways, and equipment) and outside (street noise) the classroom can cause stress in students, as measured by increased blood pressure, creating the inability to concentrate on cognitive tasks. The cumulative effect of excessive noise can result in decreased achievement on reading and other tests, as well as increased issues with behavior, attention, and concentration. Recommendations of the American Society of Heating, Refrigeration and Air Conditioning Engineers as well as the Acoustical Society of America, Working Group on Classroom Acoustics, suggest no more than 35 decibels (dB) background noise and 0.6–0.7 reverberation times with the teacher voice at 50–65 dB means a target signal to noise ratio of at least 15dB.

It is well known scientifically that natural light can have positive psychological and physiological effects on humans. Natural daylight stimulates hormones that regulate various body systems and mood. Brighter, continuous-spectrum light within a building can increase visibility, enhance alertness, reduce eyestrain, and improve mood. These physiological and psychology effects have been found to lead to better performance in a few recent studies.

A very small investment in first costs of building can have an enormous impact on the cost of running an organization by decreasing absenteeism due to various building-related sicknesses that may lead to better student and teacher health, satisfaction, teacher retention, and student performance.

Jeffrey A. Lackney

See also Charter Schools; Learning Style

Further Readings

Bingler, S., & Quinn, L. (2000, April). *Schools as centers of community: A citizen's guide for planning and design.* Washington, DC: U.S. Department of Education.

Brubaker, C. W. (2002). *Planning and designing schools.* New York: McGraw-Hill.

Duke, D., & Trautvetter, S. (2001). *Reducing the negative effects of large schools.* Washington, DC: National Clearinghouse on Educational Facilities.

Evans, G. W., Kliewer, W., & Martin, J. (1991). The role of the physical environment in the health and well-being of children. In H. E. Schroeder (Ed.), *New directions in health psychology assessment* (pp. 127–157). New York: Hemisphere.

Lackney, J. A. (2002, May). *Thirty-three educational design principles for schools and community learning centers.* Washington, DC: National Clearinghouse for Educational Facilities.

Nathan, J., & Febey, K. (2001). *Smaller, safer, saner, successful schools.* Minneapolis: Center for School Change, Humphrey Institute, University of Minnesota.

Sanoff, H. (2002). *Schools designed with community participation.* Washington, DC: National Clearinghouse for Educational Facilities.

Schneider, M. (2002, November). *Do school facilities affect academic outcomes?* Washington, DC: National Clearinghouse for Educational Facilities.

Schneider, T., Walker, H., & Sprague, J. (2000). *Safe school design: A handbook for educational leaders.* University of Oregon: ERIC Clearinghouse on Educational Management.

Tanner, K., & Lackney, J. (2006). *Educational architecture: Planning, designing, constructing, and managing environments for learning.* Boston: Allyn & Bacon.

SCHOOL READINESS

School readiness can best be defined as a psychological state in which the child is prepared for or ready to engage in and benefit from the formal and informal learning experiences, both academic and social, that organized education provides. Determining exactly what this preparedness entails, the extent to which it is active or passive, and where specifically it originates, however, introduces considerable complexity into the definition of readiness. For some, readiness is a function of the school environment and expectations, academic and otherwise, that schools have for children as they begin formal elementary education. For others, school readiness is determined by children's social skills and ability to interact in appropriate and meaningful ways with peers and teachers in the school environment. For yet others, school readiness is dependent on children's knowledge, including verbal ability, letter and number knowledge, and other aspects of acquired information important for learning in school. And for yet others again, school readiness is defined primarily in terms of children's cognitive and emotional self-regulation abilities; the ability to focus attention, to take turns, and to regulate behavior in various ways.

Each of these perspectives on readiness has merit and has been validated to some extent in studies of children's progress in school. Studies examining school expectations indicate that when enrolled in schools that implement strategies to introduce children and families to schooling prior to school entry, children achieve at higher levels and are better adjusted to school than are comparable children in schools without transition practices. Similarly, studies of social skills as aspects of readiness have indicated relations between children's social competence, friendship quality, relationships with peers and teachers, and children's academic progress and social-emotional adjustment to school. As well, studies of cognitive abilities such as general intelligence, knowledge of letters and numbers, and language comprehension and production indicate associations between these aspects of cognition, particularly language development, and academic achievement in the early elementary grades. And studies of cognitive and emotional self-regulation skills, including aspects of temperament, attention, and emotion regulation, have shown that children who are better regulated both prior to and following school entry achieve academically at higher levels and are perceived by parents and teachers to be well adjusted to school.

Given evidence in favor of multiple perspectives on school readiness, the construct is best understood as a complex developmental phenomenon, the product of multiple interacting systems or levels of influence. Within the systems approach to complex phenomena, characteristics of the child that predispose to high or low levels of attributes important for school readiness, such as basic knowledge or cognitive and social-emotional competence and self-regulation abilities, are the product of interacting genetic, physiological, and family variables that in turn influence and are influenced by characteristics of the early school environment and children's experiences in that environment. Accordingly, readiness is an emergent process

that must be uniquely considered for each individual child, rather than as an abstract standard or level of attainment that is universally applicable to all children.

Although the systems view provides an accurate conceptual depiction of the nature of school readiness, the complexity of the approach limits its applicability to specific efforts to promote school readiness among children at risk for early school failure. The promotion of readiness is essential if children are to benefit from a free and universal public education. At the turn of the 21st century, however, more than one half of a representative sample of U.S. kindergarten teachers reported substantial concerns about the readiness of their students. Furthermore, in a similar survey, teachers reported that they believed the primary obstacle to children's progress in kindergarten and the early elementary grades to be difficulty with paying attention, following directions, taking turns, and becoming meaningfully engaged in learning activities.

The emphasis in readiness that teachers voice reflects a concern with the psychological processes by which children are engaged in and motivated by learning experiences. The emphasis concerns the maintenance of an optimal level of arousal through which children are able to appropriately sustain and shift attention, to regulate both positive and negative emotions, and to engage in meaningful interpersonal interactions with teachers and fellow students. Significantly, it should be noted that a focus on motivation and engagement in readiness assumes that experiences in the classroom are motivating and engaging. Meeting this assumption, the priority of all early elementary classrooms, allows for a focus on processes internal to the child that can, to some extent, encompass emphases on social-emotional and cognitive abilities as aspects of school readiness. In theory, children who exhibit effective self-regulation are more likely to engage in purposeful social interaction with teachers and peers, acquire and apply information conveyed in learning situations, and develop a sense of self as an efficacious and successful student. Such a propitious cycle of regulation, interaction, knowledge acquisition, and sense of agency or efficacy as a student can perhaps serve as a schematic for the type of experience for which children should be ready and should and frequently do encounter in the early elementary grades.

A focus on readiness as self-regulation also offers the possibility for a definition of the phenomenon that allows for the recognition of influences ranging from the neurobiological to the neighborhood and community. Neurobiologically speaking, the age at which children typically enter formal schooling is one in which there is considerable change in the neural basis for self-regulation. Areas of the brain associated with complex reasoning and with emotional arousal, emotion regulation, and motivation are developing rapidly and in interaction during the preschool and early school years, and as such they are influenced by the types of experiences that children encounter in social and learning environments. Logically speaking, of course, school entry occurs at a time in which children are, on average, developmentally ready for the expectations of formal schooling—for sustained interactions with others on specific topics and activities that require the application of diverse cognitive and emotion regulation skills and abilities. Such developmental readiness, however, can be capitalized on only if schools provide the types of environments and experiences that support children's developing self-regulation abilities and nascent attempts at engaged and purposeful learning. In particular, such developmental readiness can perhaps be capitalized on only if children have adequate opportunity for self-direction in early school experience, including the opportunity for adequate amounts of play, including rough-and-tumble play. Given the multidimensional nature of school readiness, any attempt to define the phenomenon narrowly and to promote it through the implementation of specific strategies in isolation from other influences, such as a focus on academic but not social aspects of readiness, is unlikely to be successful.

Clancy Blair

See also Individual Differences; Learning; Maturation

Further Readings

Blair, C. (2002). School readiness: Integrating cognition and emotion in a neurobiological conceptualization of child functioning at school entry. *American Psychologist*, *57*, 111–127.

Panksepp, J. (1998). Attention deficit hyperactivity disorders, psychostimulants, and intolerance of childhood playfulness: A tragedy in the making? *Current Directions in Psychological Science*, *7*, 91–98.

Pianta, R., Cox, M., & Snow, K. (Eds.). (2007). *School readiness and the transition to kindergarten in the era of accountability.* Baltimore, MD: Brookes.

Pianta, R., & Rimm-Kaufman, S. (2006). The social ecology of the transition to school: Classrooms, families, and children. In K. McCartney & D. Phillips (Eds.), *Blackwell handbook of early childhood development* (pp. 490–507). Malden, MA: Blackwell.

SCHOOL RESOURCES

School resources are inputs into the education process, such as staff, buildings, and materials. This entry examines the definition of resources, their allocation and use in schools, and the relationship between resource use and outcomes.

Resources and Their Role in Learning

Real resources are the human and physical inputs used in education; monetary resources are the finances used to purchase real resources. A common indicator of national effort in providing education is the proportion of gross domestic product (GDP) spent on schools. In 2002, on average, Organization for Economic Co-operation and Development (OECD) countries spent 3.8% of GDP on primary, secondary, and postsecondary nontertiary institutions. This ranged from a maximum of 5.7% in Iceland to 2.6% in Turkey. There is a tendency for poorer OECD countries to spend a smaller percentage of GDP on schools than richer countries. The average for 19 OECD partner countries (18 middle-income developing countries plus Russia) was very similar in 2002 at 3.9%, but the range was much wider, from 1.2% in Indonesia to 8.7% in Jamaica. Commonly used, real, comparative resource measures are pupil-staff ratios, average class size, and students per computer.

Resources are used to provide teaching and learning for pupils and therefore are judged in relation to pupil outcomes. These outputs or outcomes of schooling are pupils' cognitive attainment measured by tests and examinations, pupils' progress to further levels of education, and earnings over their working life attributable to their education. In addition, school education can produce nonmonetary benefits for individuals and society, such as better health, better parenting, reduced crime, more active participation in community life and democratic politics, and greater social cohesion. Some of these benefits are private, as they accrue to the individual alone, whereas others are public benefits accruing to society as a whole. Thus, resources are analyzed as inputs into the schooling process and studied as part of an input-output system.

The relationship between real resources and school outputs is mediated by school context and processes of social interaction within schools. Contextual variables include the policy and governance framework within which schools operate, the locality and community the school serves, and the characteristics of its student body. Pupil characteristics, both inherent and those acquired through family upbringing and any previous schooling, determine a larger proportion of a pupil's cognitive attainment than the effects of the school itself, which also depend on the characteristics of the peer group. Thus, an additional and very important resource is the school's pupils as well as the characteristics—both individual and collective—that they bring to the school. When schools' outputs are measured after controlling for pupils' prior attainment and/or family characteristics and school context, this measure of output is known as "value added."

Criteria for Assessing Resources

Four main criteria are applied in assessing resources—adequacy, transparency, equity, and efficiency.

An adequate amount is sufficient resources to provide a specific standard of education. One way of assessing adequacy is to determine if per-pupil funding in one area is less than the average in other areas. More specifically, adequate funding is defined as the amount required to achieve defined standards for children with specific learning needs, because those with greater learning needs—often, but not exclusively, associated with social disadvantage—require more resources to achieve a given standard.

Transparency requires publicly accessible information on the resources each school receives and how it uses these resources for educating specific groups and individuals. Absence of transparency is associated with lack of accountability for the use of public funds and also with corruption, a particular problem in many developing countries. The World Bank has promoted Public Expenditure Tracking Surveys, which track the flow of resources from central government through the administrative layers down to school level. This reveals the extent to which corruption and mismanagement divert resources intended for public schools into private uses.

Equity refers to the fairness, as judged by impartial observers, with which resources in education are allocated and used. Horizontal equity is the equal resourcing of pupils with similar characteristics or learning

needs, whereas vertical equity refers to the unequal funding of students according to differences in their needs. For example, vertical equity requires students with learning difficulties to have disproportionately more resources devoted to them.

Efficiency is the relationship between inputs and outputs and can be assessed only by comparing schools. A school is internally efficient in comparison with other schools if, for a given cost, it produces the highest feasible educational attainment for its pupils. To obtain an accurate and fair measure of efficiency, inputs must include pupils' prior attainment and/or family characteristics. Internal efficiency compares the inputs used by schools with measures of output, but does not attempt to assess the social value of schools' outputs. In contrast, external efficiency is the relationship between the social value of school outputs and input costs. Given that school outputs in the public sector are not sold in markets and hence do not have market prices, there is no direct measure of their value, apart from the increase in earnings of schools' graduates that can be attributed to their schooling and not to other factors, such as ability or family background.

Resource Allocation

Resources are allocated to schools by either the public or the private sector. Public funding is mainly for state schools, but it is also provided for private-sector schools, often those with religious affiliations. Private schools may be funded by charities, but most raise the majority of their finances by fees. In developed countries, public schools are free to parents, although countries vary in charging for uniforms, equipment, and textbooks. In some developing countries, where public funds are inadequate, parents are charged fees for public schools. Clearly, the extent to which parents have to supply the resources required by schools affects demand for school places and makes it difficult for poorer parents to afford schooling. Hence, the provision of free public schooling of reasonable quality is vital for ensuring access for all social groups. In countries where weak states cannot secure adequate tax revenues or run moderately efficient school systems so as to provide universal access to schools of reasonable quality, it is argued—controversially—that private provision can fill the gap for low-income children.

In OECD countries, on average, 18% of total expenditure on schools was raised by the private sector, varying from zero in Sweden to 68% in South Korea—which Programme for International Student Assessment data show is a country with relatively high educational attainment equity. In the 19 partner countries, an average of 28% of spending was provided by the private sector, ranging from 95% in Indonesia to 10% in Malaysia.

Resource allocation to public schools varies from being highly centralized to largely decentralized, with a movement toward greater decentralization in the past 20 years. In a highly centralized system, a central or regional government allocates resources in kind to schools. The government appoints staff, provides physical resources such as books and equipment, directly pays for utilities, and maintains school buildings. In a decentralized system, the school is allocated financial resources as a lump sum budget, usually determined by a formula in which the number of students is a major, but not sole, element.

The school principal or school council, advised by the principal, is responsible for allocating the school budget for educational purposes. The school decides on its staffing establishment; appoints and dismisses staff; and pays for all building operation costs, learning resources, and other services, such as educational psychology. Decentralized school finance systems are likely to be more efficient than centralized ones. Per-student funding gives schools incentives to improve quality in order to attract parents. Resource utilization decisions at the school level provide incentives for schools to save on unnecessary expenditures and switch money into resources that will better serve the educational needs of their students, about which schools have superior information than more distant education authorities. However, social equity can be undermined by schools skimming off more able students, leading to greater social stratification between schools with reduced access to higher-quality schools for socially disadvantaged pupils.

Resource Utilization

Internal efficiency depends on selecting the best mix of resources—that is, for a given expenditure, selecting the combination of teachers, nonteaching staff, and learning resources that achieves the highest outcomes for pupils. Private schools and public schools with school-based financial management are able to make resource mix decisions. Otherwise, they are made for schools by an education authority. Monopoly power on the supply side of the labor market

constrains the achievement of efficient resource mixes in schools. Such constraints can be due to teachers having lifetime tenure as civil servants; to political power, whereby teachers are a more influential clientele for politicians than parents; or to powerful teacher unions. In such situations, teachers are likely to have inadequate incentives to work efficiently, whereas schools are constrained in their ability to divest themselves of poorly performing teachers or excess numbers of teachers. Using evidence from studies in developing countries, Pritchett and Filmer found that marginal increases in student attainment due to increases in nonstaff resources were considerably higher than those due to extra teachers; hence, there was an overuse of teachers relative to learning resources.

The average percentage of total current expenditure on staff was similar in OECD and the 19 partner countries—81% and 84%, respectively. These data give a rough indication that an 80:20 spending split between staff and nonstaff items is appropriate.

The Input-Output Relationship in Schools

Research over the past 40 years on education production functions (the relationship between school inputs and outputs) has not yet resolved whether additional school resources increase pupil attainment. A major reason it is difficult to obtain robust statistical estimates of this relationship is that most data come from natural settings rather than social experiments where the amount of resources is varied independently of the characteristics of schools and pupils. When using data from natural settings, the quantity of school resources is likely to be partly determined by the quality of the students so that there is causation in both directions—from resources to student outcomes and from student outcomes to school resources. This endogeneity in resources may occur either because governments practice compensatory funding so that lower-attaining students have more per capita spent on them or because socially advantaged parents with higher-attaining children choose better resourced schools.

There is no general consensus on the interpretation of evidence on the effects of additional resources on attainment. Levačić has noted that the available evidence for developed countries indicates that additional marginal resources allocated to a general reduction in

class sizes or pupil–teacher ratios has, at best, a small positive effect in some countries. However, additional resources targeted at particular subjects, such as math and science, and particularly at less able or more socially disadvantaged pupils, are likely to be more effective than a general increase in spending given current levels. There is limited evidence that school autonomy, if combined with external examinations, has a favorable impact on pupil attainment. For developing countries, there is clearer evidence that additional material resources, such as blackboards and textbooks, improve attainment.

In developed countries, creating incentives for more efficient use of the existing quantity of resources and targeting them more effectively at raising the performance of currently low-attaining students appear to be the most promising resourcing policies. In developing countries, increasing the overall volume of resources remains essential for expanding access and raising quality, provided that bureaucratic inefficiency and corruption can be curtailed.

Rosalind Levačić

See also Cultural Diversity; Culture; Learning; School Design

Further Readings

Levačić, R. (Ed.). (2007). The relationship between school resources and pupil attainment. In T. Townsend (Ed.), *International handbook on school effectiveness and improvement* (pp. 395–410). Dordrecht: Springer.

OECD. (2005). *Education trends in perspective: Analysis of the World Education Indicators.* Paris: UNESCO Institute of Statistics.

Pritchett, L., & Filmer, D. (1999). What education production functions really show: A positive theory of education expenditures. *Economics of Education Review*, *18*, 223–239.

SCHOOL VIOLENCE AND DISRUPTION

Concerns over school violence in the United States date from the late 1970s, with the issuance of the 1978 report, *Violent Schools—Safe Schools: The Safe School Study Report to Congress.* That report found that (a) about 25% of U.S. schools experienced vandalism, (b) 1.3% of students were victims of personal

attack, (c) 11% of students were victims of theft, and (d) 8% of secondary students missed a day of school per month out of fear for personal safety. Widespread dissemination of these statistics led to increased public and policymaker focus on school violence. A subsequent federal report in 1984, *Disorder in Our Public Schools*, cited continuing problems with school disorder nationally, leading to the passage of the Safe and Drug-Free Schools Act of 1986. By the early 1990s, school violence in the United States was perceived to have reached epidemic proportions. National attention was riveted on increasingly disturbing statistics in the early 1990s showing rising school violence, and seemingly endless school shootings throughout the 1990s, culminating with the Columbine High School shootings of April 1999. These developments, as well as events in the first 6 years of the new century, may be considered in terms of a complex nexus of issues and influences.

This entry begins with an overview of data on school violence from the early 1990s through the early 2000s. Issues of data collection and measurement are discussed with primary attention to national-level data. Dimensions of school violence and disruption (e.g., bullying) are examined, followed by a summary of research on school violence prevention. The entry concludes with a discussion of key controversies and ongoing challenges to the field.

Trends in School Violence and Disruption

There has been a pronounced decline in several measures of school violence and disruption from the mid-1990s through the early 2000s according to several measures, including data from the National Crime Victimization Survey (see Figure 1).

Some indicators suggest that a new, albeit lower level of violence has been reached. Other measures suggest a lesser decline over the years. The Centers for Disease Control and Prevention (CDC) Youth Risk Behavior Survey (YRBS) reports a significant reduction from 1993 to 2005 in students fighting in school and bringing weapons to school (see Figure 2).

A slightly different picture of stable, or slightly increasing, trends emerges, looking at the percentage of students who reported being threatened or injured with a weapon at school, and missing school due to safety concerns (see Figure 3).

In sum, the data suggest that school violence and disruption diminished over the period of 1993–2005, but that serious problems remain.

Violent Crime: At and Away From School

Year	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
At School	48	59	56	50	43	40	43	33	26	28	24	28	22
Away From School	71	70	69	58	55	59	48	39	34	28	26	32	21
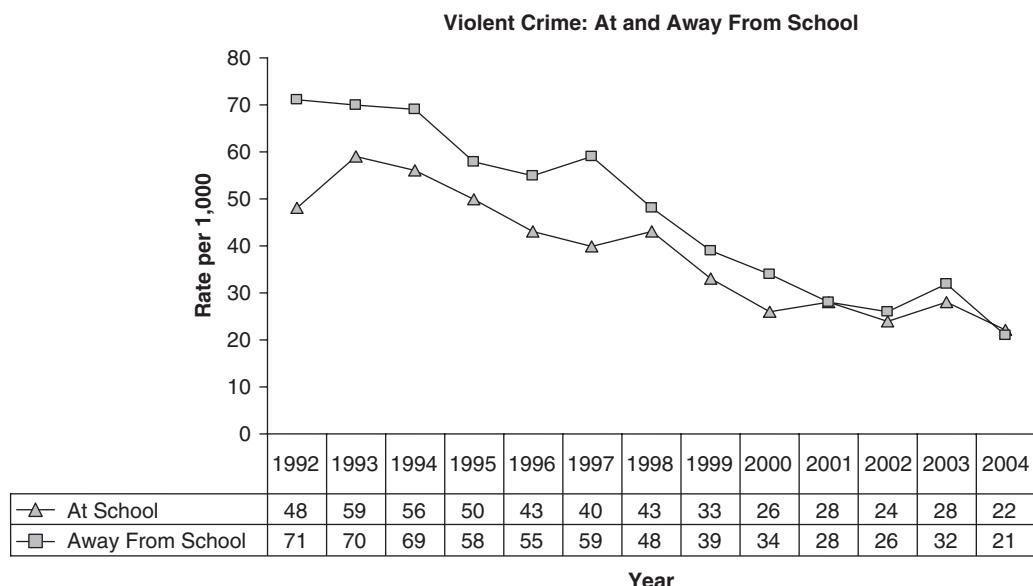
Figure 1   Rate of Student-Reported Violent Crimes Against Students at and Away From School

Source: Indicators of School Crime and Safety 2006, National Center for Education Statistics.

Year	1993	1995	1997	1999	2001	2003	2005
In Fight at School (12 mos.)	16.2	15.5	14.8	14.2	12.5	12.8	13.6
Weapon to School (30 days)	11.8	9.8	8.5	6.9	6.4	6.1	6.5
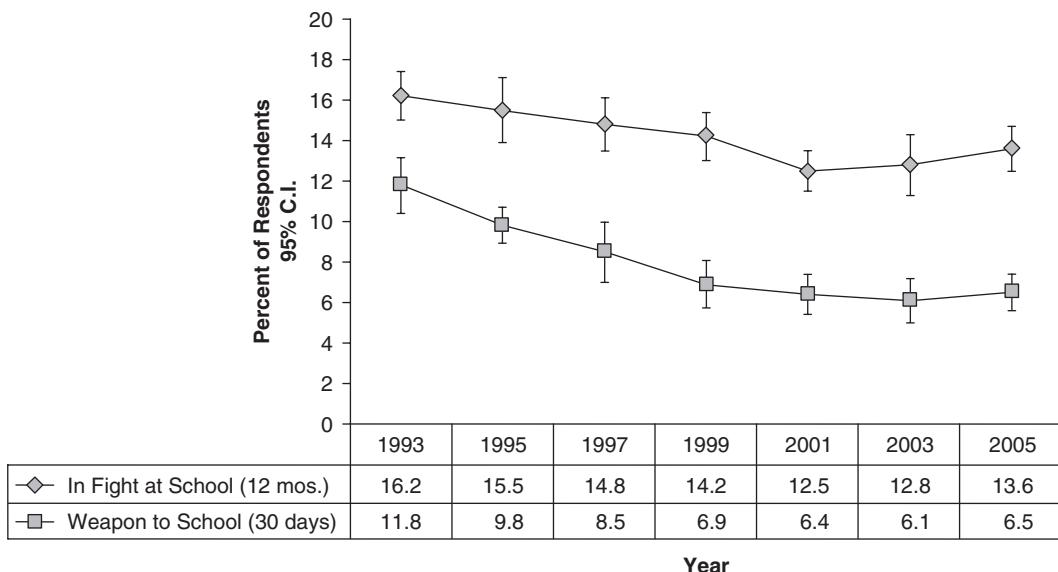
Figure 2   Percentage of Respondents Reporting Fights at School or Weapons Brought to School

Source: Youth Risk Behavior Survey (1993–2005): Fights at School; Weapons Brought to School. Centers for Disease Control and Prevention.

Year	1993	1995	1997	1999	2001	2003	2005
Threatened or Injured With Weapon at School (12 mos.)	7.3	8.4	7.4	7.7	8.9	9.2	7.9
Missed School for Safety Concerns (past 30 days)	4.4	4.5	4	5.2	6.6	5.4	6.0
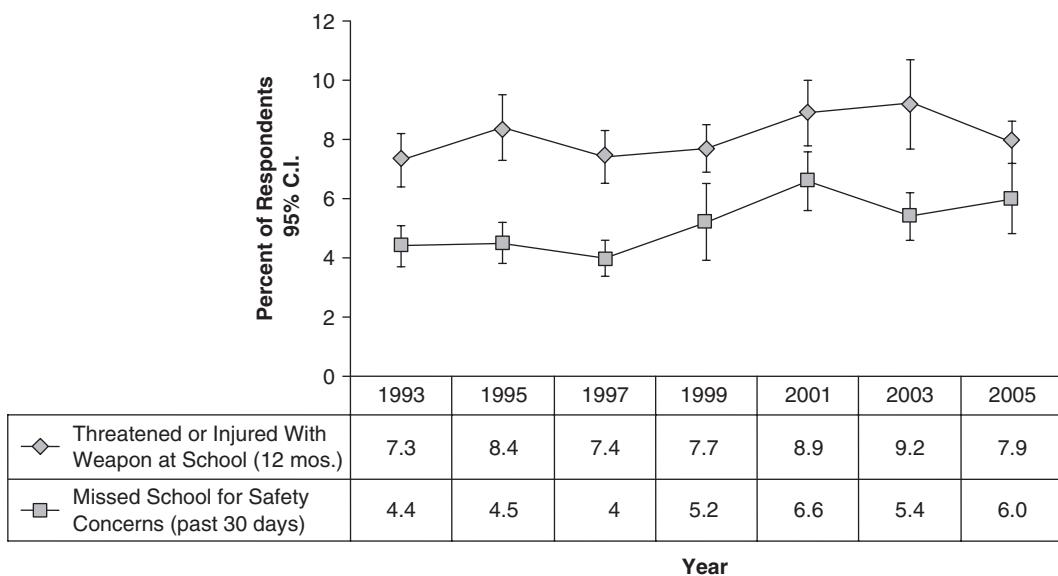
Figure 3   Percentage of Respondents Reporting Being Threatened/Injured With a Weapon or Having Missed School Due to Safety Concerns

Source: Youth Risk Behavior Survey (1993–2005): Threatened/Injured With Weapon; Missed School Due to Safety Concerns. Centers for Disease Control and Prevention.

Data Collection and Analysis

Data collection and analysis addressing school violence is challenging. Multiple agencies (e.g., Federal Bureau of Investigation, or FBI; CDC; U.S. Department of Education) collect and disseminate school violence data using agency-specific definitions and measurement tools. Some data address victimization experiences, whereas other data look at commission of crimes. The FBI's Uniform Crime Reporting Program gathers criminal activity data from state agencies based on arrest counts, which can include innocent persons and may result in biased indicators. Self-report data used in victimization surveys are prone to errors with sampling frame, instrumentation, poor respondent comprehension and recall, and telescoping effects, as well as difficulties with nonresponse. National-level data collection instruments have unique foci, design, and implementation, resulting in differing results based on technical definitions, time frames, and specific questionnaire content. Despite differences in data collection, analysis, and reporting, school violence data are informative when linked to knowledge of school violence processes. In the next section, several core dimensions of school violence are discussed.

Dimensions of School Violence and Disruption

School violence and disruption can be conceptualized as a function of their frequency, intensity, duration, persons affected, and overall impact on the school environment. Although there has been considerable national attention to highly publicized school shootings, and the public developed a perception of schools as dangerous during the 1990s, data on serious violent crime show that schools are relatively safe compared to surrounding communities. Students have less than a one in a million chance of being shot at school, an event less likely than being struck by lightning. School violence and disruption can be viewed relative to overall student behaviors at school. Figure 4 provides a conceptual depiction of
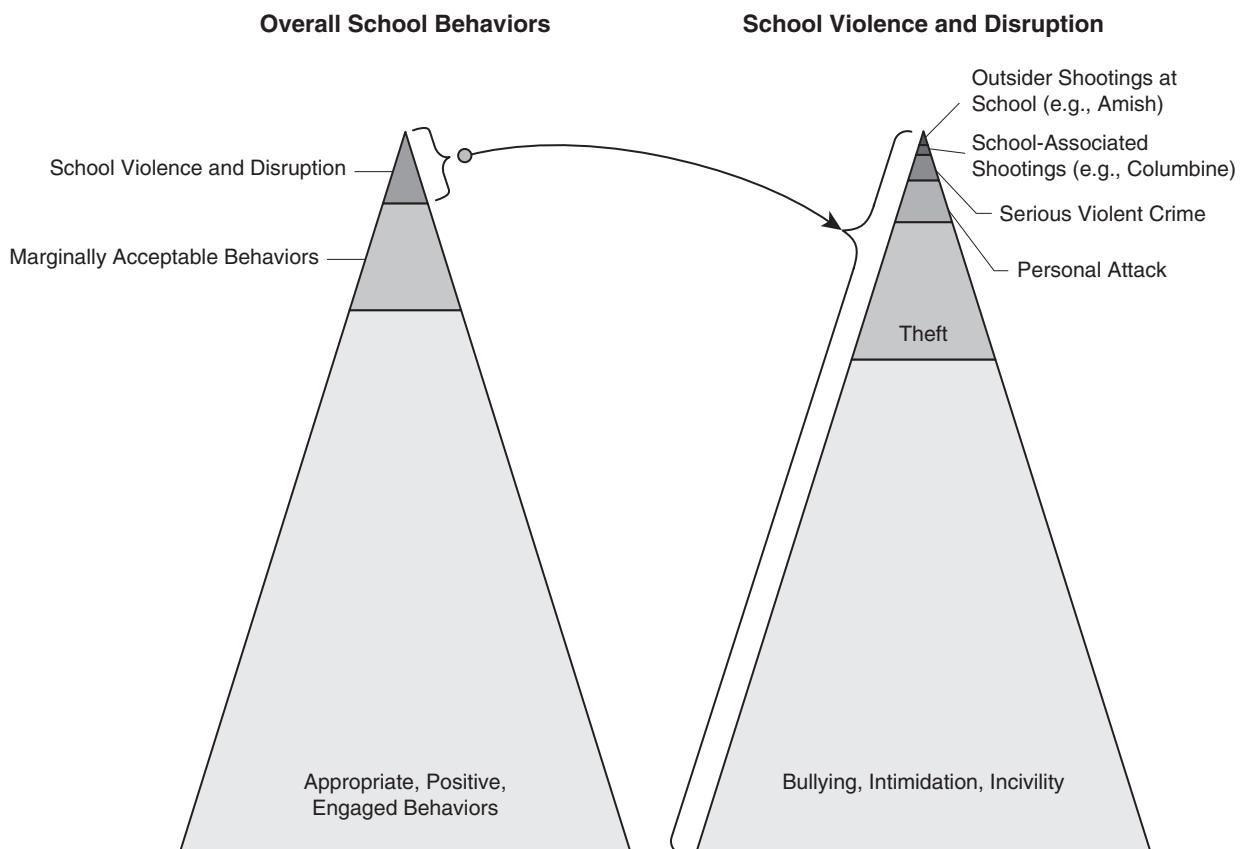
Figure 4   Conceptual Representation of School Violence as a Component of Overall Student Behaviors

school violence, clearly showing the relatively small proportion of overall school behavior it comprises. This conceptual framework provides an avenue for considering multiple dimensions of school violence and disruption.

Bullying, intimidation, and incivility collectively represent much of the antisocial behavior that occurs on a daily basis in schools. Research has focused more on bullying than on more broad indicators of intimidation and incivility. Bullying is generally defined as a function of three characteristics: (a) intentional behaviors toward another that are harmful physically, socially, or emotionally; (b) a relationship where an imbalance of power exists; and (c) repeated negative interactions. Data from the federal report *Indicators of School Crimes and Safety, 2006*, showed that in 2005, 28% of middle and high school students reported being bullied at school during the previous 6 months. Students in earlier grades are generally more prone to being bullied, and the most common forms of bullying are (a) being made fun of, insulted, or called names; (b) being the subject of rumors; and (c) being pushed, shoved, tripped, or spit on. Approximately one fourth of those reporting being bullied were injured as a result of the bullying experience, where males were almost twice as likely to be injured as compared to females. These more recent data parallel earlier findings in a 2001 study of a nationally representative sample published in the *Journal of the American Medical Association*, where approximately 30% of students were involved in bullying as bullies, victims, or both. That research found that those students who were most lonely, according to self-report, were more than 33 times more likely to be bullied, compared to students who were least lonely. Several research reports have documented near-term and longer term social and psychological harm to bully victims, including difficulties making and keeping friends and increased loneliness while in school, lower levels of self-esteem, and increased likelihood of depression as young adults. Bullies demonstrated poorer school adjustment with concomitant academic problems, were more likely to engage in drinking and smoking while in school, and were more likely to engage in criminal behavior as young adults. These data as a whole suggest the need for continued prevention programming in schools to address bullying. Related research on intimidation and incivility, including sexual harassment at school, has demonstrated linkage to psychosocial adjustment problems, including fear and anxiety, and avoidant behaviors, also indicating a need for continuing prevention measures.

Theft is the major form of school-based crime. There were approximately 55 million preschool–12 students during the 2004–2005 school year in the United States. According to National Center for Education Statistics (NCES) data, there were approximately 1.4 million school-based crimes, of which about 863,000 (62%), or 33 per 1,000 students, were theft and 583,000 (42%), or 22 per 1,000 students, were violent crime (simple assault and serious violent crime). Approximately 107,000 cases, or 4 per 1,000 students, involved serious violent crimes (rape, sexual assault, robbery, and aggravated assault).

Despite theft being more common, students' concerns focus more on fear of attack. Data from the 1999 *Metropolitan Life Survey of the American Teacher* report indicate that approximately 15% of students in public schools were worried about being victims of physical attack at or near school. More recent data from the NCES report *Indicators of School Crime and Safety 2006* mirror this finding, but to a lesser degree, showing that in 2005, about 6% of secondary students feared attack at school, about half of the 12% figure found in the 1995 NCES survey. Fear among minority secondary students was greater than for Whites, with approximately 9% of African American and 10% of Hispanic students fearing attack at school. Students' fear translates into actions, including missing school and avoiding places in and around school perceived as dangerous. For example, in 2005, according to the CDC Youth Risk Behavior Survey, approximately 6% of students in Grades 9–12 missed school at least once during the 30 days prior to the survey because of fear for their personal safety. This represented an increase from a figure of 4.4% in 1993.

Much attention has focused on school shootings over the past decade. Despite public perception to the contrary, school shootings are relatively rare events. According to NCES data, during the 2004–2005 school year, nationally, there were 21 school-related homicides. Although varying from year to year, during the early to mid-1990s, there were 28–34 school shooting deaths per year, compared to 11–21 per year during the early 2000s. The United States witnessed a dramatic change in the national experience with school shootings during the fall of 2006, where in two separate incidents, adults perpetrated shootings of students at schools in episodes that were not school-related in origin. These varied experiences with school violence

have led prevention researchers to pursue multiple avenues of effective prevention responses to school violence and disruption.

Research on School Violence and Disruption Prevention

General Research Syntheses on Prevention

Three major research syntheses on school-based violence prevention, published between 2001 and 2006, provide broad yet informative guidance. Julie Mytton and colleagues published a Cochrane Collaboration review in 2006 of a wide variety of school-based interventions to reduce student violence at the secondary level, finding an overall effect size of –0.41 for prevention programs to reduce anger, aggression, violence, bullying, and interpersonal conflict. This effect size translates to the average treatment group member (at the 50th percentile of that group) being better off than 66% of the control group as a result of the prevention program.

In 2001, a more focused meta-analysis by David Wilson and colleagues of 165 research studies of school-based interventions to reduce substance abuse, dropout rates, and conduct problems reported that non-cognitive-behavioral approaches resulted in negative outcomes or no demonstrable effectiveness. However, cognitive-behavioral and behavioral modification programs using modeling and rehearsal demonstrated low to moderate effect sizes of 0.12 to 0.37.

In another important meta-analysis in 2003, David Wilson and colleagues Mark Lipsey and James Derzon analyzed 221 studies of school-based anger reduction interventions. Effect sizes for demonstration and routine practice programs were 0.25 and 0.10, respectively. Academic, behavioral, and counseling programs showed the highest effect sizes. In sum, these syntheses suggest that violence prevention programming works, but that effect sizes are modest and varied, and that more needs to be understood as to what types of interventions best address specific needs for different groups in varied school-based contexts. More focused violence prevention research has addressed areas of threat assessment, controlling the physical environment, and whole-school approaches.

Threat Assessment

As a result of the perceived epidemic of school violence in the early 1990s and the pattern of school shootings throughout that decade, there was increased national attention on developing improved methods of school-based threat assessment, especially profiling. Early work by the U.S. Secret Service, the U.S. Department of Education, and researchers in the field demonstrated that profiling approaches designed to predict individual perpetrators of deadly violence do not work and instead inappropriately target a relatively large number of innocent individuals as suspects. Two other approaches to threat assessment are mental health assessment and automated decision-making approaches based on actuarial, artificial intelligence, and expert systems theory. Neither of these approaches has been developed to the point of demonstrating effectiveness. Dewey Cornell has developed protocols for school-based threat assessment that show promise, based on preliminary field test research using 35 schools. That approach uses a systematic, decision-tree process, where trained teams gather and evaluate an array of threat-related data.

Another approach to threat assessment, articulated by Peter Leone and Matthew Mayer, differed from the preceding methods by focusing on overall school characteristics rather than individual students. Leone and Mayer described an *unhealthy schools syndrome*, conducive to school violence and disruption, where the fit between students and school was poor. Five targeted areas contributing to poor fit were (1) academic missions that did not meet at-risk students' needs and marginalized large numbers of students; (2) heavy reliance on zero tolerance and generally punitive approaches to discipline; (3) entrenched, noncollaborative systems of control that excluded students, staff, and families; (4) racial and cultural disconnects; and (5) disconnects with students with disabilities.

Controlling the Physical Environment

Efforts to control the physical school environment have centered on equipment- and personnel-based approaches. Equipment-related approaches typically have involved using technology such as security cameras and metal detectors. There is little effectiveness research on such measures. For example, Sandia National Labs, under the sponsorship of the U.S. Department of Justice, published a report in 1999 on appropriate use of security technology in schools, but the report contained no documentation of research supporting the recommendations. There has been little or no extant research on security cameras in schools over

the past two decades. The limited body of research has focused on metal detectors. Two studies from the 1990s demonstrated that handheld and walk-through metal detectors could lower the number of weapons brought into schools, but such measures did little to change levels of violence and disruption in schools.

A larger body of research exists for the most prevalent personnel-based measure: School Resource Officers (SRO). SROs are typically certified law enforcement personnel assigned to schools with three general areas of responsibility: (a) law enforcement, (b) law-related counseling, and (c) law-related education. There have been 10 noteworthy studies of SRO programs during the period from 1995–2006, of which one was national in scope, and the others were local and regional. One of the 10 studies employed moderately rigorous quantitative methodology, examining outcomes related to program characteristics, whereas the others primarily reported descriptive data from survey administrations. Although no study to date has demonstrated a causal linkage between SRO programming and reduction of school violence, the majority of studies have reported increased parent and school staff satisfaction as a result of SRO programs in local schools, with mixed reactions from students. Several studies have reported on reductions of school violence and disruption, concurrent with ongoing implementation of SRO programs, but these results have been correlational. In summary, research on physical and personnel-based measures to manage the school environment has been relatively sparse, and most studies have been methodologically limited.

Whole-School Approaches

The most common whole-school prevention approaches are (a) security and discipline-oriented approaches, as previously examined; (b) broader intervention programs comprising multiple integrated components (e.g., Second Step); (c) systems-change programs (e.g., Positive Behavior Supports); and (d) comprehensive programming (e.g., Safe and Responsive Schools). A brief discussion of exemplar programs mentioned will illustrate the nature of these approaches.

Schools typically adopt multiple intervention programs, some mandated by local school districts, and others selected by individual school governance bodies. Second Step is a well-researched, highly effective violence prevention program for students in pre-K to ninth grade. The core components of Second Step address areas of empathy, impulse control, problem solving, and anger management, with a supplemental, family-based component available. More than 27,000 U.S. schools used Second Step during 2006.

Positive Behavioral Supports (PBS) uses a school-wide approach that systemically transforms the school environment to foster student success in academics and social interactions. PBS takes a balanced approach, developing whole-school behavioral expectations, proactively teaching and prompting students to employ prosocial behaviors, while also providing the necessary structure and discipline to respond to behavioral infractions. PBS uses a team-based approach where decision making is based on ongoing data collection and analysis. Furthermore, PBS uses a three-tier system of primary, secondary, and tertiary intervention based on research in public health. Multiple studies have demonstrated significant improvements in academic and behavioral performance of students in PBS schools.

The Safe and Responsive Schools Program (SRS) uses a comprehensive prevention framework based on three foundational elements: (1) creating a safe and responsive school climate that promotes student belonging and bonding to school; (2) early identification and intervention for at-risk students; and (3) effective responses to disruption and crisis that include policies and procedures that address more serious disciplinary infractions and crises, but that use alternatives to suspension and expulsion that help maintain a more positive trajectory of student engagement and improvement. Preliminary research has shown significant reductions in school suspensions and high levels of stakeholder satisfaction based on implementation of SRS.

Key Controversies

With a perceived epidemic in school violence in the early 1990s, and the repeated school shootings of the mid- to latter 1990s, culminating in the Columbine High School massacre in 1999, educational stakeholders focused significant attention on security measures and school discipline, including zero tolerance. Several of these approaches remain controversial, especially zero tolerance and restrictive security methods.

Zero tolerance is used to describe a variety of school-based responses to discipline. It refers to discipline methods that employ strict and inflexible punishments for transgressors with no consideration of situational factors. Zero tolerance originated in the

1980s as part of U.S. military personnel drug policies and related antidrug legislation. Schools nationwide began adopting zero tolerance policies in the early 1990s, with increased public concern about rising juvenile crime and school violence. School-based zero tolerance originally addressed the most serious behaviors (e.g., deadly weapons in schools), but in later years, it addressed a broad array of less serious student behavior. Out-of-school suspension is the primary sanction used by schools, with expulsion in the most serious instances. Advocates of zero tolerance argue that it is a necessary tool for combating school violence, where many contributing factors are beyond the control of the schools. Critics of zero tolerance policies have noted that zero tolerance does not focus on methods for improving student behaviors, nor does it respond to causes of inappropriate misbehavior. Zero tolerance has also been implicated in contributing to a downward spiral of academic failure and developing delinquency. Multiple research studies have shown that African American students have been disproportionately penalized by zero tolerance policies. There are little or no data to support the effectiveness of zero tolerance policies, and none of the five national panels on school violence has supported zero tolerance as an efficacious intervention.

Restrictive security approaches refer to a set of policies and procedures that relies predominantly on physical control and containment measures in schools, where there is relatively little attention to proactive and preventive programming. Multiple research reports have suggested that schools that rely primarily on metal detectors, security cameras, locked entrances, searches of student lockers, and security personnel regularly patrolling hallways don't reduce classroom violence. Causal research demonstrating beneficial effects of these methods is extremely rare. Researchers have repeatedly argued that schools whose programs depend almost entirely on school security measures may alienate students, making schools seem jail-like, and lead to increased behavioral problems. Recent statistical analysis of the School Crime Supplement data sets from 1995 through 2003 by Matthew Mayer demonstrate that in recent years, there has been a very limited relationship between secure building strategies and levels of school violence and disorder. In sum, although such security measures remain popular, there is a limited evidence base supporting their use.

As a follow-up to the Columbine High School tragedy in Colorado in 1999, the report of the Governor's Columbine Review Commission addressed the use of security equipment, as evidenced in this excerpt:

> Although security devices can effectively deter certain forms of school crimes, including theft, graffiti, and gang violence, they have not yet been proven to be cost-effective in preventing major school violence like that experienced at Columbine High School. Therefore, the Commission does not recommend the universal installation of metal detectors, video surveillance cameras and other security equipment as a means of forestalling school violence generally; for the present, such security devices can serve only to offer transient solutions to specific problems at individual schools. (p. xviii)

Addressing Future Challenges

Although school violence and disruption have declined over the past decade, they remain a serious concern. Research has shown that school violence reduction depends on the availability of well-coordinated, evidence-based programs that address physical safety, educational practices, and the social-emotional-behavioral needs of students. Studies demonstrate that specific supports should be provided to address mental health needs, bully prevention, and anger management/conflict resolution using a three-tier approach of primary, secondary, and tertiary prevention.

Educational stakeholders have varying knowledge, perceptions, concerns, and agendas regarding school violence, and thus, local school communities sometimes find themselves in disagreement as to the optimal course of action with regard to school violence prevention. The increased national focus on educational accountability, as exemplified by the 2002 No Child Left Behind (NCLB) law, signaled a shift toward retooling schools to focus more on demonstrating academic results. As such, many schools have had to choose between devoting time and resources to support attainment of NCLB goals, or focusing on programming and supports that may address problems with school violence and disruption.

Although there are no quick fixes to the nation's problems with school violence, the knowledge base supporting school violence prevention has been refined over the past decade, and initiatives on multiple levels seem to have helped to curb the once-growing tide of violence and disruption in schools. Experts agree that school violence is in part a reflection of deeper

societal problems, and that long-term solutions require fundamental changes across society. Continued investment in research to better understand ways to minimize school violence and disruption, coupled with broad stakeholder commitment to invest in evidence-based prevention approaches, appear to provide the best options for managing school violence in the immediate future.

Matthew Jared Mayer

See also Maturation; School Design

Further Readings

Columbine Review Commission. (2001). *Report of Governor Bill Owen's Columbine Review Commission.* Denver: State of Colorado.

Dinkes, R., Cataldi, E. F., Kena, G., & Baum, K. (2006). *Indicators of school crime and safety: 2006* (NCES 2007–003/NCJ 214262). Washington, DC: U.S. Government Printing Office.

Elliott, D. S., Hamburg, B. A., & Williams, K. R. (Eds.). (1998). *Violence in American schools.* New York: Cambridge University Press.

Jimerson, S. R., & Furlong, M. J. (Eds.). (2006). *Handbook of school violence and school safety: From research to practice.* Mahwah, NJ: Lawrence Erlbaum.

Leone, P. E., Mayer, M. J., Malmgren, K., & Meisel, S. M. (2000). School violence and disruption: Rhetoric, reality, and reasonable balance. *Focus on Exceptional Children, 33*, 1–20.

Peterson, R. L., Larson, J., & Skiba, R. (2001). School violence prevention: Current status and policy recommendations. *Law & Policy, 23*(3), 345–371.

Skiba, R. J. (2000, August). *Zero tolerance, zero sense: An analysis of school disciplinary practice* (Policy Research Report No. SRS2). Bloomington: Indiana Education Policy Center.

Scientific Method

Many methods are employed to generate and synthesize knowledge. As such, no single series of steps could contain all the strategies that can be used to discover and explain the universe. Thus, science is better viewed as a process than as a recipe or formula that can be used to create knowledge. For example, considerable knowledge is generated and passed along via careful observations and subsequent descriptions.

Basic observations and descriptions contribute to the discovery and documentation of patterns, and each of these steps is fundamental to increasing our understanding. However, the process known as the *scientific method* involves steps beyond observation and description. Specifically, these steps include development and testing of explanations for patterns derived from observations.

Discovery and documentation of patterns rooted in careful observations and descriptions certainly represents science. However, the hypothetico-deductive method is not employed when science terminates with observation and description. As such, discovery and documentation, rooted in observations and descriptions, do not comprise the process commonly termed the *scientific method.* The "scientific method" depends on the generation and testing of scientific hypotheses, which are defined as *candidate explanations.*

Scientific hypotheses typically are derived from patterns because demonstration of a pattern often generates the question: "What process causes that pattern?" The response to the question is a scientific hypothesis, or a candidate explanation. Providing a definitive answer to the question requires formulation and subsequent testing of potential explanations for the observed patterns; that is, it involves hypothesis testing.

A hypothesis generated by a pattern should not be confused with a prediction; a prediction is a statement that is likely to be factual. Most predictions can be evaluated via observation. Although the instruments used for evaluating predictions are variable and may be quite sophisticated, they should not be confused with hypotheses. The scientific method, as typically described, requires generation and testing of hypotheses. In contrast, generation and evaluation of predictions, which are fundamental to the generation of knowledge, are not part of the classically defined "scientific method." Indeed, if generation and evaluation of predictions represent application of the scientific method, then the scientific method is not unique to science. Rather, it is used for everyday activities such as mowing the lawn (the grass is too tall in some spots, so I must have missed those spots), shopping for groceries (I need carrots, so I will look in the produce section rather than searching the entire store), and commuting to work (length of route and traffic patterns dictate my path). In other words, science has little to offer beyond everyday activities if observation and description are the only means by which we acquire knowledge.

Consider a relatively simple example of a prediction: Retention of information about human behavior is greater with inquiry-based methods than with lectures. This prediction can be (and has been) evaluated with observations of students in classrooms. It also can be (and has been) evaluated with well-designed and rigorously applied experiments (involving, for example, random assignment of students to groups that are taught material in different ways). Statistical hypotheses likely will be generated and evaluated, but we should take care to distinguish between these statistical hypotheses and scientific hypotheses. Whether researchers employ observations or experiments, however, evaluation of the prediction does not involve candidate explanations. Generation of such explanations—that is, scientific hypotheses—is a logical next step in the conduct of research because it addresses the question of "why."

The ability to distinguish between predictions and hypotheses, and between statistical hypotheses and scientific hypotheses, allows us to describe a set of steps that has become known as the scientific method. First, observations are made. These observations may be formalized and synthesized in the form of general statements of patterns. The patterns often are evaluated for generality, and different patterns are determined to coincide with different events of phenomena. A typical question that emerges from patterns, especially patterns that vary with different phenomena, is "why?" The explanations—the answers to the "why" question—are hypotheses. Testing these hypotheses, typically by carefully controlling several variables and allowing one or a few variables to vary in an experiment, represents an attempt to generate knowledge. This series of steps, from observations to testing of hypotheses, commonly is termed the "scientific method."

Guy R. McPherson

See also Descriptive Statistics; Experimental Design; Inferential Statistics; Statistical Significance

Further Readings

McPherson, G. R. (2001). Teaching and learning the scientific method. *American Biology Teacher*, *63*, 242–245.

Medawar, P. (1984). *Pluto's republic*. New York: Oxford University Press.

Popper, K. R. (1968). *The logic of scientific discovery*. New York: Harper and Row.

Self-Determination

The term *self-determination* is a motivational concept that has been explicated primarily in relation to self-determination theory (SDT), proposed by Edward Deci and Richard Ryan. To be self-determined is to endorse one's actions with a full sense of choice and volition. When self-determined, individuals experience a sense of freedom to do what is interesting, personally important, and vitalizing; they experience themselves as self-regulating agents of their own behavior. Thus, self-determination signifies the experience of choice and endorsement of the actions in which one is engaged. *Autonomy* and *autonomous motivation* are terms that are frequently associated with the concept of self-determination. Although autonomy is often equated to independence in the Western world, that is not the meaning it carries within SDT. Rather, it means concurring with one's actions and feeling a sense of willingness and volition. So, within the framework of SDT, people could be autonomous even when doing something for someone else, if they fully endorsed doing so.

Autonomy is one of the three basic and innate psychological needs proposed within SDT (the other two being competence and relatedness). As such, it has been found to be a valid need cross-culturally and to be understood similarly by individuals in countries as diverse as South Korea, Japan, Russia, and the United States. When a behavior is not autonomous, it is said to be controlled. To the degree that people are controlled, they are essentially subjugated by external or internal forces that are pressuring them to behave, think, or feel in particular ways. In contrast to autonomous behaviors, controlled behaviors are experienced by individuals as coming from outside their sense of self. These behaviors are experienced as alien to the self and controlled by external demands or internal pressures.

Control Versus Autonomy

In defining what self-determination is, it seems very useful to also explore both what it is and what it is not in order to highlight its main components. As mentioned, self-determination involves being autonomous, fully endorsing one's actions, and doing what one wants to do because the activity is interesting and pleasurable (like flying a kite) or because it is

perceived as valuable (doing an important work project). Being self-determined is associated with a sense of vitality, energy, happiness, well-being, and enjoyment. When self-determined, people often report feelings of flow while engaged in the activity and display the capacity to stay involved and engaged with it. Self-determination contrasts with control, with feeling pressured, coerced, or seduced. Self-determination is not doing something because someone else makes you, nor is it doing something to get someone else's approval. It is not doing something in order to get a prize, a reward, or some other external incentive, for those are all examples of being controlled. Being controlled or non-self-determined is associated with feelings of pressure, whether it comes from external demands or from internal threats of guilt.

The total absence of self-determination is embodied in what is called *amotivation*, which represents a total lack of motivation for an activity. As part of SDT, Ryan and Deci proposed various types of motivation that can be aligned along a continuum of increasing self-determination or perceived autonomy in performing an activity. The various forms of motivation are presented from left to right in increasing degree of self-determination, from the least self-determined (amotivation) to the most self-determined (intrinsic motivation). External regulation and introjection are considered controlled forms of motivation and therefore not self-determined. In contrast, identification, integration, and intrinsic motivation are considered to be autonomous forms of motivation and therefore self-determined. This entry considers each of these concepts in turn.

Amotivation

When amotivated, people are alienated and lack the intention to do a behavior. They feel a complete lack of choice and volition, and often, they will not behave at all. If they do behave, it is without the intention to do so and they may not know why they are behaving. Amotivation is associated with disengagement and often with negative affect and self-disparagement. Individuals who are amotivated would endorse statements such as "I'm not sure why I'm taking this class anymore. I think that maybe I should quit."

Controlled Forms of Motivation

External regulation and introjection are the two forms of controlled motivation on the continuum. *External regulation* is the first type of extrinsic motivation and is the type that is commonly referred to as extrinsic motivation—that is, it is behavior done to get rewards or avoid punishments. In general, extrinsic motivation signifies that a behavior is not done for its own sake but instead is a means to an end, and the external-regulation type of extrinsic motivation involves people feeling pressured and obligated to engage in the activities because of external contingencies. External regulation is not self-determined, and individuals who perform activities out of external regulation do not feel a sense of volition while engaging in these activities. For example, individuals who are externally regulated would agree with statements such as "I am doing my homework because I have to do it to pass the course."

External regulation is maintained by the presence of external contingencies, but people can internalize these contingencies and become somewhat more self-determined for the corresponding behaviors. Deci and Ryan argue that internalization and integration of aspects of one's environment are natural human processes. With respect to extrinsic motivation, they are the means through which external contingencies can become the basis for more autonomous regulation of extrinsically motivated behaviors. Human beings are naturally inclined to strive for higher levels of self-determination, and internalization and integration are the processes that actualize this tendency. Then, as the external contingencies become increasingly internalized within one's self, they lead to different forms of regulation that are increasingly self-determined.

Introjection is a form of motivation in which the external contingencies maintaining a behavior have been internalized, but only partially so. Basically, people take in a contingency or regulation without really accepting it as their own, so the contingency is controlling them even though the regulation is now within them. So, instead of being maintained by external pressures or rewards, introjected behaviors are now maintained by internal pressures such as ego involvement or the avoidance of feelings of guilt. Behaviors regulated by introjects are still experienced as controlled and thus are not self-determined. For example, individuals who are behaving out of introjected regulation would agree with statements such as "I am taking this class because I would feel bad about myself if I didn't." Although the level of self-determination a person experiences when he or she is behaving out of introjection is low, it is still greater

than the level experienced when one is externally regulated.

Autonomous Forms of Motivation

Identification, integration, and intrinsic motivation are the three autonomous or self-determined forms of motivation. *Identification* is the first form of autonomous motivation on the continuum of self-determination. When behaving out of identification, individuals feel volitional because they have accepted the regulation as their own. Therefore, the activity is performed because it is personally valuable and important to the individual, and the person has a sense of endorsing his or her actions. For example, individuals who are behaving out of identification would agree with statements such as "I am reading this assignment because it helps me to develop skills that are important to me." Although identification is still a form of extrinsic motivation, it is nonetheless relatively self-determined because it is engaged in with a sense of choice and volition.

Integration is an even fuller form of self-determined extrinsic motivation. The activity is still a means to an end—for example, people do it because it is instrumental for attaining some valuable, self-selected goal—but because they have integrated the identification with other aspects of themselves, they do the behavior with a feeling of self-determination. Integrated regulation is the most self-determined form of extrinsic motivation. When an activity is regulated through integration, it is experienced as an important part of the individual's identity. For example, individuals who are behaving out of integration would agree with statements such as "I am studying these art history texts because learning about such things is part of who I am."

Intrinsic motivation represents the prototype of self-determination. When intrinsically motivated to do something, individuals engage in the behavior out of pure interest and for personal satisfaction. The activity is performed solely because it brings pleasure and satisfaction to the individuals while they are doing it. The activity is an end in itself rather than being a means to a goal or outcome. Because intrinsic motivation is based in people's interest in the activity rather than in its instrumental value, intrinsic motivation is inherent to the individuals and does not result from internalization of extrinsic motivation. Individuals who are behaving out of intrinsic motivation would agree with statements such as "I am majoring in physics because I really find it interesting and enjoyable."

Outcomes Associated With Self-Determination

An abundance of empirical research has supported the importance of self-determination for human growth, development, and well-being. Social contexts that support individuals' natural tendencies toward active engagement and psychological growth will serve to enhance their self-determination. For example, in situations where people are provided with choices and options, offered a rationale for engaging in certain behaviors, and supported in their pursuits, greater feelings of self-determination are reported. In contrast, situations that are pressuring and controlling, where few choices are provided, tend to lead to lower levels of self-determination. In sum, autonomy-supportive situations—that is, those that offer choices, options, and supports—lead to higher levels of self-determination and in turn to positive outcomes, whereas controlling situations lead to lower levels of self-determination and more negative outcomes.

According to Ryan and Deci, the natural tendencies for healthy development and optimal functioning are defined using the concept of basic psychological needs, which are defined to be innate, universal, and essential for health, well-being, and growth. The basic psychological needs are a natural aspect of human beings, and therefore, they apply to all people regardless of gender, group, or culture. To the extent that individuals' needs are continually satisfied by social contexts, people will experience self-determination, function effectively, have satisfying relationships, and develop in healthy ways. However, to the extent that individuals' needs are thwarted by social contexts, people will experience lower levels of self-determination and show evidence of ill-being, poor relationships, and nonoptimal functioning.

Chantal Levesque, Kelly J. Copeland,
and Edward L. Deci

See also Home Environment and Academic Intrinsic Motivation; Intrinsic Versus Extrinsic Motivation; Motivation; Motivation and Emotion

Further Readings

Chirkov, V., Ryan, R. M., Kim, Y., & Kaplan, U. (2003). Differentiating autonomy from individualism and

independence: A self-determination theory perspective on internalization of cultural orientations and well-being. *Journal of Personality and Social Psychology*, *84*, 97–110.

Deci, E. L., Eghrari, H., Patrick, B. C., & Leone, D. R. (1994). Facilitating internalization: The self-determination theory perspective. *Journal of Personality*, *62*, 119–142.

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, *125*, 627–668.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*, 227–268.

Levesque, C., Zuehlke, A. N., Stanek, L. R., & Ryan, R. M. (2004). Autonomy and competence in German and American university students: A comparative study based on self-determination theory. *Journal of Educational Psychology*, *96*, 68–84.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*, 68–78.

Ryan, R. M., & Deci, E. L. (2006). Self-regulation and the problem of human autonomy: Does psychology need choice, self-determination, and will? *Journal of Personality*, *74*, 1557–1585.

SELF-EFFICACY

Over the past several decades, educators and educational psychologists have been interested in students' self-perceptions. This concerted attention to students' views of themselves is quite understandable. From the perspective of an educator, enhancing the positive self-perceptions of learners is an important educational goal. Educators desire students to become autonomous individuals who have strong and positive feelings about themselves as they engage in the myriad responsibilities and demands of school and life. Educational psychologists too have been interested in student self-perception, particularly as such perceptions may affect student motivation, persistence, and ultimately, classroom learning. The goal of the following discussion is to examine research and theory relating to learners' perceptions of self-efficacy.

Self-efficacy is a central construct in Albert Bandura's social-cognitive learning theory. It is defined as an expectation that one holds regarding one's capabilities to accomplish a particular task or goal. In a most general and perhaps somewhat rough sense, self-efficacy might be thought of as a sense of confidence to act successfully. Individuals who believe that they can overcome obstacles and achieve have what is often called by social-cognitive learning theorists a sense of personal agency. Self-efficacy is an important part of one's feelings of personal agency and a crucial element of agentic views of motivated learning.

Although self-efficacy may at first glance be similar to notions of self-concept and self-esteem, it should be emphasized from the onset that it is quite separate from these other, broad self-perceptions. Self-efficacy is an expectation that is related to a given task. As such, it is a specific self-perception that is directed toward accomplishing a goal. Self-efficacy cannot be understood nor assessed without a task as the context of one's self-appraisal. In contrast, self-concept refers to one's collective self-perceptions across many different tasks and assembled from many different interactions. To be sure, self-concept is multidimensional, but it is referenced to broad domains, such as academic self-concept or physical self-concept. Self-esteem is also different from self-efficacy. As usually defined, self-esteem is a self-perception that refers to the judgment of worth that one places upon one's self. Such judgments of worth rely heavily on social norms, social comparisons, and individual values; they do not rely on one's task-related appraisals of one's capabilities to execute the requirements of a task. Rather, self-esteem focuses on self-valuing.

Self-efficacy can be distinguished further from other kinds of expectations that have been the focus of psychologists who are interested in student motivation. In particular, self-efficacy is distinct from the psychological notion of outcome expectations. These expectations concern students' beliefs regarding the consequences of achieving tasks. For example, students may fully expect a reward or other outcome when completing a task, but they may have grave doubts about their capacities to accomplish the task that would earn them such a reward. In other words, expectations of the consequences of behaviors are conceptually distinct from self-beliefs concerning enacting behaviors. The former is an outcome expectation; the latter is an efficacy expectation.

Self-Efficacy and School Learning

A very large body of research literature demonstrates that self-efficacy is an important predictor of human

performance across a vast array of tasks and settings. Although this section is devoted to the relationship between self-efficacy and learning in schools, it is noteworthy that self-efficacy is of wide theoretical importance and practical application. Indeed, research has examined the relationship between self-efficacy and human change in clinical, counseling, sports, health, and other settings. The bulk of these studies has shown that self-efficacy is a moderate predictor of performance across many different behaviors. These findings are particularly valuable because they speak to the generalizability of the self-efficacy construct and hence its widespread usefulness in understanding human change.

Turning to the school context, self-efficacy is a moderate predictor of academic achievement. Students with higher self-efficacy tend to achieve more. Furthermore, heightened levels of self-efficacy are associated with longer persistence on classroom tasks. The relationship between self-efficacy and performance is a complex one. Self-efficacy may reflect prior accomplishments, such as how one has done in the past on a particular kind of arithmetic problem, but it is not the same as simply knowing what one can do. Research in a number of different settings has demonstrated that self-efficacy predicts future performance even when past performance is controlled. Stated somewhat differently, self-efficacy provides information about student achievement that is over and above information provided by students' past accomplishments. It is more than simply knowing what one has done previously. It is a motivational belief concerning personal agency that energizes willful execution of knowledge and skill when students are confronted with academic challenge.

Several studies have concluded that self-efficacy is not just correlated with academic achievement, but that it mediates academic learning. Using complex mathematical models, these studies have suggested that students' prior achievements enhance self-efficacy, which in turn affects subsequent achievement. In this way, the relationship between self-efficacy and achievement is a reciprocal one. As learners master academic tasks, they experience heightened self-efficacy. In turn, this heightened self-efficacy enhances subsequent learning. Thus, self-efficacy and learning are mutually enhancing as both intertwine across time.

Enhancing Student Self-Efficacy

With the importance of student self-efficacy well established, researchers have turned their attention to examining how students gain a sense of self-efficacy. Research and theory point to four main sources of efficacy information. First, students can gain efficacy beliefs by directly experiencing the mastery of classroom tasks. As students engage and successfully accomplish academic goals, they are imbued with ever-increasing beliefs of self-efficacy. Self-efficacy gleaned from mastery experiences is robust and reliable. Success expectations build directly on repeated successful experience.

A student may also witness the mastery demonstrations of others as they are engaged in academic work. Through a process of observational or vicarious learning, students can experience enhanced perceptions that they too can successfully complete tasks. As efficacy beliefs formed in this way are indirect, that is, rely on the observations of others' task performances, they are necessarily less reliable sources of efficacy information. After all, a student must identify with the skill level and abilities of fellow students before adopting a similar level of self-efficacy.

A third source of efficacy information is found in the persuasive verbal comments of others. Students often hear the encouraging words of teachers and parents who exhort them to persist in meeting academic challenges. These statements of persuasion are used frequently to bolster student self-efficacy. Once again, however, enhancing student self-efficacy by exhortative means is not as predictable as changes in self-efficacy that arise from students' direct mastery experiences. Students may discount the remarks made by supportive teachers and parents, particularly if such remarks are overdone.

Finally, students' efficacy beliefs may be affected by their physiological state, that is, their perceptions of physiological agitation, heart rate, breathing rate, and so on. A student who feels agitated prior to and during a test, for example, may be instilled with a sense of dread and lowered self-efficacy. Students who interpret their physiological arousal in such a way would be predicted to have lower motivation and achievement. On the other hand, another student may experience the same level of physiological arousal when confronted with a testing situation and interpret such arousal as normal. Indeed, this student may interpret test jitters as a harbinger of appropriate arousal. The important observation here is that physiological information may enhance or detract from efficacious beliefs depending on one's cognitive interpretation of the physiological state.

Teachers can play a crucial role in setting learning conditions that can promote self-efficacy in their students. Because mastery experiences are a dominant source of students' efficacy beliefs, teachers who carefully set mastery experiences in the classroom will produce students with a confident sense of their own capabilities. One way in which teachers can enhance student efficacy is to provide learners with moderately challenging tasks. When students overcome moderate challenge, they experience heightened efficacy expectations. On the other hand, students who accomplish easy tasks would not be expected to experience heightened self-efficacy because such experiences would only confirm prior capabilities. Similarly, students who are confronted with very difficult tasks would likely experience failure and a resultant loss of efficacy.

Learners who take ownership of their achievement in the classroom are able to experience heightened self-efficacy. Not all learners, however, take such ownership. Students who have experienced repeated failure may adopt a sense of learned helplessness. Even when successful, these students may attribute their successes to factors outside of their control, such as work being easy, teachers being lenient, or simply being lucky. Teachers can prevent and remediate such disabling thinking by providing students with feedback that highlights that they are the origin of their successes. When teachers indicate to students that their successes are the result of their capabilities, teachers communicate a way of thinking to students that alerts them to take ownership of task mastery.

Mastery experiences can be made more salient when learners are instructed to set goals for themselves. Goal setting helps to initiate students' self-appraisals as learners monitor their progress toward meeting clear endpoints of their efforts. Several studies have noted that students who set small, short-term goals rather than one larger, long-term goal are more motivated and experience heightened self-efficacy. The reason for this is easy to see. Setting a series of small and more proximal goals permits more opportunities for accomplishment. More frequent successes build more confirmation of growing expertise and higher levels of self-efficacy.

As mentioned previously, vicarious experiences can boost self-efficacy. Witnessing the skilled demonstrations of fellow students can imbue observers with an increased sense that they too can engage in the demonstrated actions. Teachers can improve learning and motivation gleaned vicariously by using a number of strategies. For example, research indicates that teachers who use multiple models to demonstrate skills will increase the transmission of the demonstrated behavior and enhance self-efficacy.

Much more might be said about enhancing student self-efficacy. However, the foregoing shows the importance of putting motivational beliefs on the agenda of classroom teaching. Students learn a sense of their capabilities to tackle classroom work, and teachers play a crucial role in such learning. Teachers can build instructional experiences that produce students who are confident, persistent, and active in their engagement of classroom tasks.

Future Promise of Self-Efficacy Research

Self-efficacy has become a central construct in understanding human change and its promotion. With that said, several areas have received relatively less attention. The means of promoting student self-efficacy is becoming well understood, however, less well understood is how teachers' sense of instructional efficacy might be enhanced. Teachers, like the learners they teach, complete classroom work. Although their work is different from that of their students, it too is composed of numerous tasks completed with various self-beliefs regarding their perceived competence. How teachers develop instructional efficacy, what supports teacher efficacy, and how teachers' self-efficacy influences student learning warrant further explication.

A second and particularly exciting area of self-efficacy research concerns the notion of collective efficacy. Social organizations in general and schools in particular might be thought of as having a collective sense of confidence to act. This kind of efficacy may be present when groups adopt goals, bring forth plans and strategies, and judge a collective sense of community empowerment. How collective school efficacy might influence both student and teacher self-efficacy has not been established clearly. Such an analysis would require a complex, multilevel approach that would assess the unique and shared effects of collective, teacher, and student levels of self-efficacy on learning.

Finally, studies of the developmental trajectories of students' efficacy beliefs should be conducted. In particular, the factors that affect long-term development of these self-beliefs across schooling would provide

a broad view of how children and adolescents acquire personal agency.

John Walsh

See also Learned Helplessness; Motivation; Self-Determination; Self-Esteem; Social Learning Theory

Further Readings

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: Freeman.

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, *52*, 1–26.

Bong, M., & Clark, R. (1999). Comparison of self-concept and self-efficacy in academic motivation research. *Educational Psychologist*, *34*, 139–154.

Pajares, F. (1996). Self-efficacy beliefs in achievement settings. *Review of Educational Research*, *66*, 543–579.

Tschannen-Moran, M., Hoy, A. W., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research*, *69*, 202–248.

SELF-ESTEEM

Self-esteem refers to one's overall positive or negative evaluation of oneself and can be viewed in a variety of ways: as an enduring trait, as stable versus fragile, as specific domains of contingent self-worth, or as a goal that people pursue. Self-esteem is thought to develop in childhood and remains relatively stable over time. Although there are costs associated with different types of self-esteem, there are alternatives to pursuing self-esteem that may help to reduce the costs.

There is tremendous emphasis in North American culture on self-esteem, with continuing discussion over whether self-esteem is a universal need or is unique to North American culture. In popular culture, thousands of self-help books and child-rearing manuals have been written on this topic; in the scientific community, more than 15,000 journal articles have been published in this area in the past 30 years. Self-esteem has often been hailed as an antidote to society's problems. The self-esteem movement of the 1990s assumed that raising people's self-esteem could reduce social problems such as low academic achievement, high dropout rates, teenage pregnancy, domestic violence, and drug and alcohol abuse. A recent review of the literature by Roy Baumeister and colleagues, however, suggests that the effects of trait self-esteem may be limited. High self-esteem, once thought to predict a variety of positive outcomes, was found to be associated only with a tendency to experience positive emotions and feelings of self-efficacy. Low self-esteem, once thought to be a cause of negative outcomes, was found to be merely a symptom or correlate of negative outcomes rather than a reliable cause. Overall, their review suggested that there may be more to self-esteem than whether it is high or low.

This entry examines the multifaceted nature of self-esteem: how self-esteem affects people's responses to self-threats; the stability of self-esteem; contingencies of self-worth, or the domains on which people base their self-esteem; and self-validation goals, or the desire to prove that one possesses certain qualities on which self-esteem is based. This entry concludes with a discussion of cross-cultural differences in self-esteem and ways to reduce the costs associated with pursuing self-esteem.

Responses to Self-Esteem Threats

High-self-esteem people possess favorable, confidently held self-views. They are skilled at maintaining and enhancing their self-esteem, especially in the face of self-threats. For example, following failure feedback, high-self-esteem people, compared to low-self-esteem people, are likely to call to mind their strengths relative to weaknesses, persist longer on tasks, take more risks, and affirm themselves in domains unrelated to the threat in order to boost their self-esteem. In addition, they are likely to deflect the threat away from themselves and make self-serving attributions, such as by blaming others or dismissing the validity of the negative feedback.

Low-self-esteem people possess relatively less favorable, less confident self-views and are highly concerned about the possibility of being rejected by others. When low-self-esteem people fail, they are quicker to call to mind their weaknesses relative to strengths, and tend to internalize and overgeneralize the negative feedback to themselves, viewing the failure as confirmation of their global inadequacy. Low-self-esteem people engage in self-protective strategies following failure to avoid further loss of self-esteem, such as reducing their task motivation, avoiding risks, and showing decreased desire to repair their negative moods.

Low-self-esteem people possess contingencies of interpersonal acceptance; when they fail, they automatically associate failure with being rejected by

others. Indeed, low self-esteem is related to personality measures reflecting insecurity in relationships, such as rejection sensitivity, insecure attachment styles, and excessive reassurance-seeking in relationships. Following failure, low-self-esteem people become more interpersonally oriented, which leads them to be liked more by strangers. In interactions with strangers, low-self-esteem people may be more highly attuned to the possibility of rejection and thus likely to expend much effort to ensure that the other person likes them. However, in the context of close relationships, low-self-esteem people show self-protective responses following self-threat; they distance themselves from their partners, devalue their relationships, and doubt their partner's love for them—preemptive responses to avoid the possibility of being rejected by their partners.

High-self-esteem people are confident of their interpersonal acceptance and do not chronically worry about whether others will reject or accept them. In fact, high-self-esteem people become more independent following failure, focusing on their unique abilities and competencies rather than on their interpersonal qualities, leading them to be liked less by strangers. However, in close relationships, high-self-esteem people draw closer to their partners and affirm their relationships following self-threat, thereby repairing their self-esteem.

Research on high- and low-self-esteem people's responses to success are mixed. Some studies show that both high- and low-self-esteem people show increased self-esteem and positive mood following success. However, other studies suggest that low-self-esteem people experience heightened anxiety following success, because positive feedback is inconsistent with their negative self-views. Indeed, low-self-esteem people are likely to dampen experiences of positive affect, whereas high-self-esteem people are likely to savor them. Such findings are consistent with William Swann's self-verification theory, which states that people are motivated to verify or confirm their existing self-views, even if they are negative.

Stability of Self-Esteem

In addition to trait self-esteem, researchers have also investigated the stability of people's self-esteem. In particular, Michael Kernis and colleagues have distinguished between stable high self-esteem and fragile high self-esteem. People with stable high self-esteem show relatively few fluctuations in feelings of self-worth over time, because their self-worth is less dependent on external feedback or circumstances. In contrast, people with fragile high self-esteem show dramatic fluctuations in their feelings of self-worth; because they depend heavily on feedback from others and the external environment to determine their self-feelings, they experience instability of self-esteem and increased anger and hostility toward others upon receiving negative feedback. Indeed, narcissists are especially likely to have fragile high self-esteem— although they think highly of themselves, they possess an underlying fragility of self-worth that leads to an insatiable quest to be continually validated and admired by others.

Contingencies of Self-Worth

Contingencies of self-worth reflect specific domains on which people base their self-worth. Jennifer Crocker and colleagues identified seven domains of contingency among college students: academic competence, physical attractiveness, others' approval, competition, virtue, having support from one's family, and having God's love. Individuals differ in their bases of self-worth, with consequences for how they spend their time, their mental and physical health, and interpersonal outcomes. For example, people who base self-worth on attractiveness spend more time shopping, grooming, and socializing; people who base self-worth on virtue spend more time volunteering and attending church. Research has shown that certain domains of contingency incur more costs to mental and physical health than others. For example, basing self-worth on appearance is associated with depressive symptoms and eating disorders, whereas basing self-worth on virtue is associated with less alcohol use and higher grade point average (GPA). Basing self-worth on academics, although associated with more time spent studying, does not predict GPA; one explanation is that staking self-worth on academics leads to increased stress and anxiety, which, in turn, interferes with optimal performance.

Although contingencies are highly motivating, they are also a source of vulnerability because people are likely to experience negative events in domains of contingency that threaten their sense of self-worth. For example, people who base self-worth on academics report more negative self-evaluative thoughts following failure than those who are less contingent on their academic competence. Furthermore, low-self-esteem people who base self-worth on academics and

experience failure are likely to disengage from the goal to appear competent and be quicker to associate themselves with failure than with success. Finally, research has shown that among college seniors applying to graduate schools, those whose self-worth was highly contingent on academics showed dramatic fluctuations in self-esteem and mood in response to acceptances versus rejections from graduate schools, which, in turn, led to instability of self-esteem and an increase in depressive symptoms over time.

Interpersonally, research has shown that high-self-esteem people who base self-worth on academic competence and receive an academic self-threat become less responsive to others' personal problems and ultimately become less likeable. In contrast, low-self-esteem people who are contingent on academics and experience academic threat become more supportive toward others' personal problems and thus are more liked by others.

Self-Validation Goals

People are likely to adopt self-validation goals aimed at proving and validating that they possess those qualities on which their self-esteem is based. Research has shown that students who base self-worth on academics are likely to possess academic self-validation goals aimed at proving their competence to others. Although self-validation goals are motivating, when students focus on validating their ability, they are likely to experience more academic anxiety, show less intrinsic motivation, make more negative ability-based attributions, and perform more poorly on tasks following failure than those who are not as focused on demonstrating their ability.

Culture and the Pursuit of Self-Esteem

Although some researchers have proposed that self-esteem is a fundamental need, others have argued that self-esteem is not universal but, instead, a cultural construction unique to cultures emphasizing individuality and personal achievement. For example, Steve Heine and colleagues have shown that in East Asian cultures such as Japan, there is more focus on self-improvement and self-criticism, whereas in North American culture, the focus is on protecting and enhancing self-esteem.

Recently, self-esteem has been conceptualized as a goal pursuit, in which people adopt the goal to protect, maintain, and enhance their self-esteem. In North American culture, people spend much of their time and energy pursuing self-esteem. This pursuit is highly motivating because when people succeed, they experience boosts to their self-esteem and increases in positive mood. However, because success is not guaranteed, people are also likely to experience threats to their self-worth, leading to drops in self-esteem and increases in negative affect. Over time, the pursuit of self-esteem may interfere with the fulfillment of psychological needs such as competence, autonomy, and relatedness with others. For example, when people are focused on protecting self-esteem, they are less open to learning from their mistakes or examining areas that they need to improve.

An alternative to pursuing self-esteem may be to shift from goals focused on protecting and enhancing self-esteem to goals that are larger than the self and include others, such as adopting learning goals aimed at making a contribution to one's field. Along these lines, research has shown that students who adopt mastery goals in the classroom show deeper cognitive processing and more intrinsic interest than those with performance-oriented goals.

Lora E. Park

See also Goals; Motivation; Self-Efficacy

Further Readings

Baumeister, R. F., Campbell, J. D., Krueger, J. I., & Vohs, K. D. (2003). Does high self-esteem cause better performance, interpersonal success, happiness, or healthier lifestyles? *Psychological Science in the Public Interest*, *4*, 1–44.

Blaine, B., & Crocker, J. (1993). Self-esteem and self-serving biases in reactions to positive and negative events: An integrative review. In R. F. Baumeister (Ed.), *Self-esteem: The puzzle of low self-regard* (pp. 55–85). New York: Plenum.

Brown, J. D., & Dutton, K. A. (1995). The thrill of victory, the complexity of defeat: Self-esteem and people's emotional reactions to success and failure. *Journal of Personality and Social Psychology*, *68*, 712–722.

Crocker, J., & Park, L. E. (2004). The costly pursuit of self-esteem. *Psychological Bulletin*, *130*, 392–414.

Crocker, J., & Wolfe, C. T. (2001). Contingencies of self-worth. *Psychological Review*, *108*, 593–623.

Heine, S. J., Lehman, D. R., Markus, H. R., & Kitayama, S. (1999). Is there a universal need for positive self-regard? *Psychological Review*, *106*, 766–795.

Kernis, M. H. (1993). The roles of stability and level of self-esteem in psychological functioning. In R. Baumeister (Ed.), *Self-esteem: The puzzle of low self-regard* (pp. 167–182). New York: Plenum.

Swann, W. B., Jr. (1996). *Self-traps: The elusive quest for higher self-esteem.* New York: W. H. Freeman.

Vohs, K. D., & Heatherton, T. F. (2001). Self-esteem and threats to self: Implications for self-construals and interpersonal perceptions. *Journal of Personality and Social Psychology*, *81*, 1103–1118.

Sex Education

Sex education may be defined narrowly as a reproductive health curriculum delivered to young people by public school teachers. This limited view can be considered problematic because it focuses only on sexual functioning or behavior. A more current term— *sexuality education*—conveys a broader meaning that includes attitudes, roles, and relationships, as well as social and cultural aspects of being female or male. Also, sexuality concepts are learned throughout life from a range of people and resources in a variety of settings. The Sexuality Information and Education Council of the United States (SIECUS) defines *sexuality education* as a lifelong process of acquiring information and forming attitudes, beliefs, and values about identity, relationships, and intimacy. Sexuality education addresses the sociocultural, biological, psychological, and spiritual dimensions of sexuality by providing information; exploring feelings, values, and attitudes; and developing communication, decision-making, and critical-thinking skills.

For sexuality education to be comprehensive, it should be included at every school grade level. Educational psychology guides the determination of age and developmental appropriateness for both content and methods. Curriculum design uses theories involving motivation, social and cognitive learning, and behavior change. In addition to developmental and theoretical foundations, programs should be sensitive and respectful toward community values and beliefs as well as the cultural background of students. SIECUS believes that parents are their children's primary sexuality educators, but programs that are school-based, originate in religious or community groups, or come from health care professionals can complement and augment the education children receive in their families.

Although the specific content and focus of sexuality education varies and may be contentious, research polls have consistently demonstrated that the vast majority of U.S. parents (87%–94%) want schools to provide education on topics including growth and development during puberty, sexual abstinence, sexually transmitted infections (STIs), HIV/AIDS, contraception, and disease prevention methods. Students have also been polled and express strong support for sexuality education that includes coverage of health risks, birth control, abortion, and how to handle sexual feelings. Whatever the specifics, the primary goal of sexuality education is the health and well-being of learners. The purpose, history, and challenges of sexuality education are addressed in this entry.

Purpose

The goal of providing balanced and accurate information is based on the premise that some sources of information are incomplete or misleading. When children between 10 and 12 years old are asked, 38% say they get "a lot" of information about sex from TV, movies, and magazines and 31% say they get "a lot" from friends. Among teens (13 to 15 years old), friends are ranked highest (64%) and media are next (61%). The Internet (40%) falls above mothers (38%) but below schools and teachers (44%) as a source of information about sex. In school-based sexuality education, factual information that emphasizes realism and the connection between sexuality and positive human relationships can counteract distortions and misrepresentations.

Reducing the harmful effects of sexual behaviors is a goal espoused by many. The targets are most often reducing unwanted pregnancy rates; limiting the spread of STIs, including HIV; restricting sexual activity (sometimes based on marital status, age, or sexual orientation); and addressing sexual harassment, assault, and child sexual abuse. It is understandable that some of these concerns make sexual health an obvious goal among public health priorities; they formed the basis of a report titled *The Surgeon General's Call to Action to Promote Sexual Health and Responsible Sexual Behavior.* However, changing behavior may be asking more from an educational intervention than is warranted. Evaluations of various sexuality education programs demonstrate changes

in knowledge and, in some cases, altered attitudes. Impacts on behavior are complicated; documentation shows that sexuality education does not increase sexual activity, and there is evidence that protection is used more consistently following certain programs. On the other hand, signing virginity pledges and hearing about condom failure rates rather than effectiveness seems to increase the likelihood of broken pledges and unprotected sexual intercourse. Efforts to define what is normative and expected sexual behavior and problems with such efforts are discussed in the ideological segment that follows.

Preparing students to become sexually healthy adults with all the skills, attitudes, and values that that entails is another goal of sexuality education. The focus on health involves moving away from an emphasis on harm, danger, and secrecy. Instead, young people are given opportunities to develop relationship insights; they can then recognize what to seek rather than what to avoid. Sexually healthy relationships are honest; consensual; nonexploitative; mutually pleasurable; safe from harm, both physical and psychological; and protected from unwanted pregnancy and STIs. Sexuality education aimed at fulfilling this purpose helps students understand their obligations and responsibilities; develop interpersonal skills including communication, decision making, and assertiveness; and form caring, supportive, and noncoercive relationships. The central purpose of sex education has not always been viewed this broadly.

History

A variety of lenses can be used to examine how the past influences the present. Inventions and research topics reflect the concerns of the time. Antisex crusaders Sylvester Graham in the early 1800s and John Harvey Kellogg in the late 1800s both introduced products (graham crackers and corn flakes) intended to improve health by calming children and preventing masturbation. Examples of varied research methods to learn about sexual behavior include Alfred Kinsey's interviews of thousands of men and women in the mid-20th century, William Masters and Virginia Johnson's physiological study of human sexual response in the second half of the 20th century, and the Centers for Disease Control and Prevention's (CDC) administration of the Youth Risk Behavior Surveillance Survey (YRBS) to students across the nation. All of these examples mirror societal views throughout history of what is important to know and respond to.

Ideological

Dennis L. Carlson frames his historical overview as four distinct ideologies, each covering about a quarter of the 20th century, but all still present in sexuality education today to varying degrees. These ideologies incorporate a relatively coherent set of values and beliefs that appears to represent the society as a whole but is based on the interests of one group or class. The first ideology is traditional and is based on an essentialist theory that sexuality must be repressed for the survival of civilization because it represents sickness and sinfulness. In the classrooms of today, the traditional ideology is represented by the "just say no" approach and characterized by slogans such as "pet your dog, not your date" and "control your urgin' and be a virgin." A central aim is to enumerate the social and emotional risks of sexual intimacy, to censor and repress sexual expression for the overall good of society.

A second ideology is progressive. It is a rational, science-based reaction in which the secular state manages sexual and social problems. The sexual ethic of this ideology is grounded in limited expression rather than repression. Both contraception and abortion serve a social function (limit population and poverty) with access under state control. This ideology can be critiqued today for its almost unquestioning support of government initiatives to solve social problems and its implicitly racist and elitist view of society.

The third ideology is called radical Freudian and is based on the recapture of pleasure and love. It is not surprising that the curriculum that represents this ideology moves far beyond issues of biological functioning. The sexual libertarianism of Herbert Marcuse is central to this approach, and creativity in work and play, for example, are reactions to an earlier emphasis on hard work and discipline.

Finally, the fourth ideology is libertarian; sexual diversity and individual sexual rights are primary. This perspective moves beyond the conventional views of sexuality as vice or virtue, normality or perversion. Individuals are the ultimate judges of their own behavior; ethical codes are based on reciprocal granting of freedom of choice and consensuality. The earlier description of a sexually healthy relationship based on mutuality, consent, and nonexploitation is an

example of sexuality education components based on this libertarian ideology. Inclusion of the range of sexual orientations and gender identities is another example of the affirmation of diversity as opposed to emphasis on conformity.

Pedagogical

Several educators and authors (Evonne Hedgepeth, Joan Helmich, and Douglas Kirby) have divided the recent history of sexuality education into generations according to the pedagogy employed. The first generation is knowledge-based; telling is the main method and the cognitive domain is the only one addressed. The underlying philosophy is, "If they know it, they will do it."

A second generation put greater emphasis on values clarification and skill building, especially communication and decision making. This approach incorporates the affective domain and leads to greater personal self-awareness and tolerance of the values of others, but impact on behavior is limited. A paraphrase of this method might be, "If they believe it is right, they will do it."

Sexuality education of the third generation arose in reaction to the first two. Specific moral messages focus on restricting sexual activity to heterosexual marriage. Providing information about prophylaxis is rejected because of the belief that it gives a double message; "If they don't know about it, they won't do it but if we tell them about it, they will do it."

The fourth generation recognizes the strengths and weaknesses of the previous methods, focuses upon behavior change and learning theories, and does affect behavior. The theme of this generation is, "If they practice what they know and apply what they believe, they will be equipped to make wise choices."

The fifth generation of sexuality education is still to come. Its development emphasizes integration of theory, research, and health promotion; it is comprehensive, positive, and proactive as opposed to reactive. As with previous generations, there are challenges to be addressed.

Challenges

If the effectiveness of sexuality education is based upon outcomes, including the rates of unwanted pregnancy, abortion, and incidence of STIs, then the United States is far down the list compared to other developed countries, especially in western Europe. The reasons are complex, but lack of support for early and ongoing sexuality education, mixed messages ranging from explicit media images to lack of access to contraceptive methods and information, limited teacher preparation, and fear of opposition all contribute.

Ethical challenges also exist. One important question is whether it is appropriate to apply a single set of beliefs or values to a diverse population. Another issue is the potential harm that could result from withholding information. A third concern involves health disparities based on limited access to services such as methods of contraception and disease prevention. Feminist scholars Lynn Phillips and Michelle Fine put forward a vital challenge. They recognize the need for courage and capacity to look critically at social arrangements, rigorously analyze power differences, and interrupt the structured silences that deprive students of the important sexuality education they need and deserve.

Sarah C. Conklin

See also Abstinence Education; Curriculum Development; Gender Identity; HIV/AIDS; Sexual Orientation

Further Readings

Advocates for Youth. (2002). *Resource guide for sex educators: Basic resources that every sex educator needs to know about.* Washington, DC: Author.

Advocates for Youth. (2003). *Science and success: Sex education and other programs that work to prevent teen pregnancy, HIV and sexually transmitted infections.* Washington, DC: Author.

Illinois Caucus for Adolescent Health. (2007). *Curriculum content review: An in-depth look at curricula in use in Illinois classrooms.* Chicago: Author.

Satcher, D. (2001). *The Surgeon General's call to action to promote sexual health and responsible sexual behavior.* Rockville, MD: Office of the Surgeon General.

Sexuality Information and Education Council of the United States. (2004). *National Guidelines Taskforce: Guidelines for comprehensive sexuality education kindergarten–12th grade* (3rd ed.). New York: Author.

Taverner, W. J., & Kelly, M. Issue 7. Should sex ed teach about abstinence? In W. J. Taverner (Ed.), *Taking sides: Clashing views in human sexuality* (10th ed., pp. 104–123). Dubuque, IA: McGraw-Hill.

Wiley, D. C. (2002). The ethics of abstinence-only and abstinence-plus sexuality education. *Journal of School Health*, 72(4), 164–167.

SEXUAL ORIENTATION

Sexual orientation refers to one's emotional, romantic, and sexual attraction to another person and is generally assumed to be heterosexual, or *straight*, meaning that one is attracted to members of the opposite sex. However, various studies suggest that between 5% and 10% of the population may be homosexual, meaning that these people are emotionally, romantically, and sexually attracted to members of the same sex. This topic is pertinent to educators, administrators, school psychologists, and others who work with children and youth because adolescents often begin to become aware of their sexual orientation before, during, and after puberty. Additionally, homosexual youth, including college-age young adults, often cope with bullying, assaults, and even murder in some cases, as well as shame, depression, and low self-esteem because of internalized negative societal views regarding homosexuality (i.e., internalized homophobia); these youth are two to three times more likely to attempt suicide than their heterosexual peers. This entry examines several topics related to sexual orientation, including common terminology, self-disclosure (i.e., coming out), possible causes, and other information.

Common Terminology: Gay, Lesbian, Bisexual, and Transgender

Although the word *gay* is often used synonymously with *homosexual, gay* generally refers specifically to gay men, whereas the word *lesbian* refers to a woman who is attracted to another woman. Use of the words *gay* and *lesbian* is generally thought to be preferable, rather than *homosexual*, the former of which are thought to be more personable, whereas the latter is considered more clinical and diagnostic, thus implying pathology to be treated and cured, although homosexuality was removed from the *Diagnostic and Statistical Manual of Mental Disorders* in 1973.

Additionally, *bisexuality* refers to both men and women who are equally attracted to members of both sexes. Alfred Kinsey and his colleagues noted that sexual orientation exists on a continuum with homosexuality at one end and heterosexuality at the other end, but that many people were in the middle (i.e., bisexual). The Kinsey Scale is well-known and used by researchers today, even though it dates back to

1948. Although the *transgender* community has unique challenges and needs, it is part of the GLBT community, a common acronym referring to gay, lesbian, bisexual, and transgender, sometimes also referred to as *queer*, a previously pejorative term reclaimed by some in the GLBT community as a word of empowerment.

Coming Out: A Process

Coming out of the closet is a common phrase that refers to the process by which a GLBT person becomes aware of his or her sexual orientation and begins to disclose it to trusted friends and family members. Vivienne Cass, an Australian psychologist, was one of the first to put forth a developmental model of homosexual identity formation, as have others. These models tend to be stage oriented and generally follow a similar linear process whereby GLBT identity begins, sometimes as early as elementary school ages, with an awareness that *something is different.* Gradually, this difference becomes understood as having to do with attraction (e.g., an adolescent boy develops a crush on another boy, rather than a girl). These attractions may be in direct conflict with societal norms, religious values, and many overt and covert messages (e.g., only "sissy" boys like other boys). Thus, many GLBT people will *go into the closet* as they become aware of their sexual orientation, hiding it, while presenting themselves as heterosexual, even, in some cases, to the point of engaging in heterosexual marriage and having children.

However, these GLBT feelings rarely go away, and so the closeted GLBT person may begin to seek information and the company of other GLBT people, often first going online, where the seeming anonymity of the Internet provides a sense of security. If and as the GLBT person becomes increasingly comfortable with the idea that he or she *may* be GLBT, possibly experiencing his or her first GLBT sexual encounter, the GLBT person will likely begin to acknowledge his or her sexual orientation, first to him- or herself, and then to a trusted friend or family member. This can have devastating results if the GLBT person is rejected by his or her family or friends, but it can also be a boost to self-esteem if the GLBT person is accepted.

It is common for GLBT people in this stage to be out of the closet in some parts of their life (e.g., with close friends or family members), but not out in other

areas (e.g., at work or school). This process can be further complicated with ethnic minority GLBT people who may be managing dual identities (e.g., African American, female, and lesbian), and it can be prematurely rushed if GLBT people are *outed* by others, either intentionally or unintentionally (i.e., disclosing a GLBT person's sexuality without permission). Gradually, however, the GLBT person may come out to more people, maybe becoming an activist for GLBT rights, and eventually reach a point where being GLBT is not seen as the end-all-be-all of his or her existence. Rather, similar to the way sexuality may be one of many aspects in a heterosexual person's life, so also a GLBT person's sexuality may be an important part of who he or she is, but the GLBT person may see him- or herself as more (e.g., a professional, amateur athlete, sibling, and published author).

Nature Versus Nurture: Causes of Sexual Orientation

Although male homosexuality has, in the past, been thought to be caused by an absent father or a domineering mother (a stereotype that has fallen out of favor), there is no known cause at this time for sexual orientation, whether homosexual or heterosexual, despite various attempts to find a biological or genetic link, as well as a number of studies that found high homosexual concordance rates in monozygotic (identical) and dizygotic (fraternal) twins. One theory, the *prenatal hormonal hypothesis*, suggests that homosexuality may be influenced by prenatal hormonal levels at critical fetal developmental stages, as observed in laboratory animals (e.g., male rats presenting themselves to other male rats in a female receptive posture after having had prenatal hormonal manipulation). Although compelling, it is unclear if the same holds true in humans. Similarly, there are a number of published studies suggesting a biological or genetic connection to sexual orientation, but to date, these studies remain inconclusive and are often criticized for methodological reasons or a lack of independent replication. Currently, it is generally held that sexual orientation, whether heterosexual or homosexual, may develop through several factors, including biological and environmental influences, as well as genetic predispositions.

It should also be noted that therapeutic attempts to change sexual orientation, commonly referred to as either *reparative therapies* (i.e., those with

a psychotherapeutic approach) or *transformational ministries* (i.e., those with a religious approach), are generally considered ineffective, changing, at most, only sexual *behavior*. Reparative therapies have been rejected by many major organizations that focus on education (e.g., American Association of School Administrators, American Counseling Association, American Federation of Teachers, American Psychiatric Association, American Psychological Association, National Association of School Psychologists) for several reasons, including (a) their outdated basis that homosexuality is a mental disorder, (b) the concern that they may be harmful to a GLBT person's self-esteem, and (c) their unproven track record.

GLBT Pride and History

Although there had been a number of early *homophile* organizations (e.g., the Daughters of Bilitis and the Mattachine Society), the modern GLBT rights movement is generally thought to have started with the Stonewall Riots in New York on June 27, 1969, coincidentally a few days after the death of Judy Garland, a GLBT icon. The Stonewall Inn was (and still is) a gay bar, and on the night of June 27, 1969, it was raided by the New York City Police Department. At that time, outward displays of homosexuality were illegal (e.g., dancing with someone of the same sex), as is still the case in many parts of the world today, punishable in some places by death. Police raids on gay bars were somewhat frequent during that time, but on that particular night, the patrons of the Stonewall Inn, some dressed in *drag* (i.e., men dressed as women) and others being homeless GLBT street youth, had had enough. They resisted arrest and fought back, throwing coins and even an unearthed parking meter at the police, who, surprised, were reported to have found themselves barricaded in the bar at one point. The riot lasted several nights; the bar was burned, but the modern GLBT rights movement had begun.

Then came AIDS. And millions of people worldwide, heterosexual and homosexual, have been and continue to be infected and affected by it. Although HIV/AIDS was first thought to be a "gay disease" because it decimated many gay men in major metropolitan areas, such as New York and San Francisco, and was even called the "gay cancer" by the popular press and GRID (gay-related immune deficiency) by the medical community in the mid-to late 1980s, it

should be noted that HIV/AIDS knows neither political boundaries nor sexual orientation. HIV/AIDS mobilized the GLBT community and brought it out of the closet in a way that nothing before ever had, and this virus, along with other sexually transmitted infections (STIs), remains a valid concern of the GLBT community. Thus, those who work with a GLBT population should be aware of this history and caution clients and students to consider either abstinence and/or safer-sex practices (e.g., talking with their partners about STIs and using condoms) when choosing to engage in sexual activity.

Discriminatory Terminology

With the burgeoning academic field of queer studies, a number of words and phrases have been coined to describe discriminatory assumptions and practices toward GLBT people, including Adrienne Rich's expression, *compulsory heterosexuality*, to describe societal assumptions and demands that people be heterosexual. Other phrases include *homophobia*, the irrational fear and/or hatred of GLBT people; *heterosexism*, the presumption that someone is heterosexual; *internalized homophobia*, the internalization by GLBT people of negative societal views and stereotypes, often contributing to low self-esteem and/or depression and suicidal ideation; and *gay lifestyle* or *homosexual lifestyle*, a phrase generally used by GLBT opponents to imply that being GLBT is a lifestyle choice of multiple sexual partners, STIs, alcohol and substance abuse, and ultimately, depression and premature death. The phrase has negative connotations, is inappropriate for those who work with a GLBT population, and fails to recognize that being GLBT, per se, does not lead to depression and its many manifestations (e.g., alcohol and substance abuse); rather, because of negative societal views toward GLBT people, internalized homophobia can lead to psychological distress and its many manifestations. Similarly, *sexual preference* also implies a choice and is not appropriate when working with GLBT people; *sexual orientation* is the preferred term.

Future Directions

Hardly a century ago, homosexuality was a moral question, mostly under the purview of religious leaders. It then became a legal issue for the criminal justice system to consider. Psychology brought it into the medical realm as a disorder to be diagnosed and treated. With each step, GLBT people have moved from "sinful" to "criminal" to "sick" to today's increasing understanding and acceptance that human sexuality, whether heterosexual, homosexual, or bisexual, is simply part of the human condition, and perhaps love, wherever it is found between consenting adults, is an extraordinary thing to be celebrated rather than condemned.

Andrew D. Reichert

See also Gender Bias; Gender Differences; Sex Education

Further Readings

American Psychological Association. (2000, February). *Guidelines for psychotherapy with lesbian, gay, and bisexual clients.* Retrieved December 10, 2006, from http://www.apa.org/pi/lgbc/guidelines.html

Byne, W. (1997). Why we cannot conclude that sexual orientation is primarily a biological phenomenon. *Journal of Homosexuality*, *34*(1), 73–80.

Byne, W., & Parsons, B. (1993). Human sexual orientation: The biologic theories reappraised. *Archives of General Psychiatry*, *50*, 228–239.

Carter, D. (2004). *Stonewall: The riots that sparked the gay revolution.* New York: St. Martin's.

Cass, V. C. (1979). Homosexual identity formation: A theoretical model. *Journal of Homosexuality*, *4*(3), 219–235.

Remafedi, G. (Ed.). (1994). *Death by denial: Studies of suicide in gay and lesbian teenagers.* Boston: Alyson.

Rich, A. (1980). Compulsory heterosexuality and lesbian existence. *Signs: Journal of Women in Culture and Society*, *5*, 631–660.

Shilts, R. (1987). *And the band played on: Politics, people, and the AIDS epidemic.* New York: St. Martin's.

SHAPING

The term *shaping* was coined by B. F. Skinner to refer to both a specific teaching technique and a general learning process. Shaping is the technique whereby a behavior totally outside the learner's current repertoire of skills is carefully brought into existence by reinforcing successive approximations to the desired performance. Shaping also refers to the learning process that takes place whenever a behavior is modified or molded over time by the consequences that differentiate its various forms.

Although not always the case, shaping is most commonly employed when, for whatever reason, the learner can neither follow verbal instructions on how to perform the desired skill nor readily imitate the behavior of another who can perform the skill. When shaping a behavior, the teacher (or therapist or trainer) starts by reinforcing whatever specific action the learner currently exhibits that is most like the new skill the teacher wants the learner to acquire. Often, the response initially selected for reinforcement bears little resemblance to the performance that is ultimately desired, even though the teacher always selects the closest approximation to that final action that the learner is now likely to perform. When repeated reinforcement has made this initial response more frequent, reinforcement of that behavior is stopped and the behavioral criterion that must be met to obtain further reinforcement is shifted to an action that is slightly closer to the ultimate performance desired. Typically, this new criterion behavior is one that, although not likely to occur at the start of training, has begun to appear during the period when responses that met the previous criterion were getting stronger, but sometimes, it is the discontinuation of reinforcement or extinction of the previously reinforced behavior that stimulates occurrence of variations that meet the new criterion. Thus, the technique takes advantage of closer approximations to the desired behavior that emerge naturally, along with other variations in behavior, during the course of reinforcing and extinguishing earlier approximations. When the newly selected criterion behavior is occurring reliably, the teacher again changes the requirement to a behavior in the direction of the final goal. Teaching proceeds in this fashion, often quite rapidly, moving smoothly across a continuum of criteria for reinforcement that approximate ever more closely the final performance desired, until that performance is attained. It is for this reason that shaping is also known as *the method of successive approximations.*

Through skillful shaping, novel, elaborate, and even peculiar acts never previously performed can be brought into existence. For example, shaping has been used to teach parrots to roller skate and ride bicycles, raccoons to play basketball, and chickens to peck out tunes on miniature pianos. Far less frivolously, shaping has also been used to teach speech to language-delayed children and to help stroke victims and other brain-damaged individuals recover functional use of affected arms and hands. As a learning process, shaping is inherent to any change in behavior where a skill is seamlessly altered or honed because of its direct and immediate consequences. Modern coaches of sports or art forms where precision in skilled movement is important (e.g., gymnastics, figure skating, ballet) take advantage of this fact by deliberately and strategically providing salient and immediate differential response consequences in their coaching practice.

Although astute teachers, coaches, and animal trainers have probably used shaping intuitively for millennia, it was not until the groundbreaking research of B. F. Skinner in the 1930s and 1940s that shaping became an explicit part of the technical and theoretical knowledge base of formal academic psychology. Skinner found that when he gradually increased the amount of force a hungry laboratory rat had to exert on a lever to obtain a food pellet reinforcer, the rat learned to press the lever harder and harder until, eventually, it pressed it far harder every time than it ever would have done otherwise. Similarly, he found that when he gradually increased the amount of time the rat had to hold down the lever before releasing it would yield a food pellet, the rat learned to hold down the lever for far longer durations than it ever would have done normally.

This process of changing a *quantitative* characteristic of a behavior (e.g., grams of response force, seconds of response duration) by gradually increasing the amount of that quantity required for reinforcement had never before been scientifically described or explored. Skinner initially called the process *the differentiation of a response.* A few years later, he discovered that *qualitative* or *topographical* features of behavior (i.e., specific postural adjustments and limb movements) could also be systematically modified via the method of successive approximations by simply watching the learner and delivering the reinforcement at the appropriate moment by hand, rather than by sensing the behavior electromechanically and presenting reinforcement via automatic control apparatus, as he had always done before. This finding came as a great flash of insight to Skinner and had a profound impact on his subsequent scholarship. It was at this point that he coined the term *shaping* to refer to the process by which the physical or social environment differentially reinforces quantitative and/or qualitative variations in behavior so as to bring about more elaborate or refined, proficient, and efficient patterns of behavior.

Just as the phenomenon of artificial selection informed Charles Darwin's view of the role of natural

selection in biological change or evolution over generations, the phenomenon of shaping informed Skinner's view of the role of operant conditioning in behavioral change over an individual's lifetime. The basic principles at work in the operant conditioning of an individual's behavioral habits and abilities over a single lifetime parallel the principles at work in the evolution of a species' biological traits and capacities over successive generations. In evolution, natural selection of slight biological variations via differential reproduction in a given generation gives rise to a different distribution of variants in the next generation upon which the selection process operates again. Over generations, this eventuates in new forms of life quite different from the ancestral forms. In operant conditioning, selection of slight variations in an individual's behavior via differential reinforcement gives rise to further variations upon which selection by consequences operates again. Over time, this eventuates in forms of behavior quite different from the earlier topographies.

In his later writings about human behavior, Skinner distinguished between *contingency-shaped behavior* and *rule-governed behavior.* The former referred to behavior that, as described above, is shaped through direct contact with its actual consequences. Both humans and nonhuman animals exhibit contingency-shaped behavior, but only people exhibit rule-governed behavior (i.e., behavior that originates from verbal instructions or rules stated by others). However, Skinner saw shaping as having at least two fundamentally important influences on rule-governed behavior. First, the habit of following directions or complying with rules comes into existence itself via shaping (i.e., via direct contact with the consequences of obeying and disobeying rules). Second, after the tendency to obey rules has become established so that new behaviors originate purely through following verbal instructions, those new behaviors are subject to further modification by the direct consequences of actually performing them, and thus the contingency-shaped behavior process again comes into play. The same thing happens to behavior that originates via imitation. Thus, shaping is an important and pervasive process among the many mechanisms that influence the particulars of human behavior.

Gail B. Peterson

See also Behavior Modification; Learning; Operant Conditioning

Further Readings

Peterson, G. B. (2004). A day of great illumination: B. F. Skinner's discovery of shaping. *Journal of the Experimental Analysis of Behavior*, *82*, 317–328.

Skinner, B. F. (1938). *The behavior of organisms: An experimental analysis.* New York: Appleton-Century-Crofts.

Skinner, B. F. (1951). How to teach animals. *Scientific American*, *185*, 26–29.

Skinner, B. F. (1953). *Science and human behavior.* New York: Macmillan.

SHORT-TERM MEMORY

Short-term memory refers to the part of the memory that has both brief duration and limited capacity. In contrast to long-term memory, which can store an infinite amount of information over a long period of time, the storage of information in short-term memory is restricted and lasts only a few seconds in the absence of rehearsal. A recent view of short-term memory, called *working memory*, suggests a more active system in which information is not only stored but also processed. Working memory is a mechanism for temporarily storing and manipulating the information to perform a variety of complex cognitive tasks. Short-term memory, or working memory, is believed to play a central role in learning, reasoning, and comprehension.

Characteristics

Short-term memory has some important characteristics. First, short-term memory is limited in duration. Information in short-term memory is kept for only a short amount of time before it is transferred to long-term memory or it decays. Generally, short-term memory retains information for only 15–30 seconds. In order to keep information longer, information must be rehearsed frequently. With rehearsal, the information will reenter the short-term memory and be held for a further period. The longer information is stored in short-term memory, the easier it is to manipulate information to perform complex cognitive tasks.

Second, short-term memory has a limited storage capacity. Basically, the capacity of short-term memory is approximately seven items, plus or minus two. The total amount of information that can be stored in short-term memory might depend on the particular

rules used to manage the information. Research has demonstrated that the chunking of information can increase the capacity of short-term memory. The process of chunking can reduce the loading of short-term memory by decreasing the number of items and increasing the size of each item. This chunking strategy allows the brain to automatically cluster certain items together and thus to remember the items much longer. This is why a hyphenated phone number is easier to remember than seven individual digits. Chunking is an important strategy for transferring short-term memory into long-term memory.

The third important feature of short-term memory is that it is subject to forgetting. Information in short-term memory is forgotten easily and quickly in the absence of further processing. Forgetting is mainly caused by decay, displacement, and interference. Decay is the primary mechanism of memory loss. Information in short-term memory decays over time. Short-term memory fades away after a few seconds without rehearsal. In addition, forgetting may occur through displacement: New information pushes out old information from short-term memory and replaces it. Moreover, interference seems to play a role in forgetting. Short-term memory is sensitive to disruption. Memory from earlier trials can interfere with recall of later ones. Information in long-term memory may also interfere with information in short-term memory.

Finally, short-term memory is regarded as a gateway to long-term memory. Attention plays an important role in selecting information to be held. Research has suggested that the perception of the importance of information determines to a great degree what information is subjected to further processing. Forgetting may begin with the information that has the least importance. Thus, the meaningfulness of an item is critical for its retention. Retention in short-term memory allows the opportunity for information to be transferred into long-term memory. Encoding and rehearsal are two main processes by which transfers occur from short-term memory to long-term memory. Encoding strategies may include *substitution*, which involves replacement of the incoming information by another symbol, or *elaboration*, which refers to developing a newly formed memory trace. Other factors, such as the intention to learn, the presence of incentives, interest in the information, and the level of arousal, also can affect encoding. Rehearsal allows more opportunity to encode into long-term memory. Activities such as using verbal rehearsal, forming mental images, and actively organizing information in short-term memory may enhance long-term retention. Adoption of rehearsal or encoding may depend on the existing knowledge in long-term memory. Information is remembered better if it can be linked to other known facts. With more existing knowledge about incoming information, there would be less effort to engage in rehearsal. However, passive residence in short-term memory may not be sufficient to transfer information into long-term memory.

Educational Applications

There are several applications of short-term memory and working memory in the field of education. Short-term memory, or working memory, is assumed to play an important role in explaining language comprehension, problem solving, and age differences in memory. The process of language comprehension takes time. The first words of a sentence need to be remembered until the end of the sentence is produced in order for the entire meaning of the sentence to be comprehended. In language comprehension, short-term memory might be critical for providing connection in reading and listening. Short-term memory may also be necessary for problem solving. Mental problem solving requires attention, encoding, storage, and manipulation of information. Several different working memory processes contribute to problem solving. Information must be encoded into working memory, other information must be retrieved from long-term memory, steps in problem solving must be arranged in order, and unrelated information must be excluded. Working memory is viewed as having a critical role in planning and organizing these types of activities, which are necessary for problem solving. Finally, working memory might be the key to understanding age differences in memory. Research has demonstrated that there is greater age-related decline in working memory. Aging is linked to slower processing, as found by slower encoding of target items into memory or slower switching from one task to another.

Shu-Chin Yen

See also Episodic Memory; Explicit Memory; Long-Term Memory; Memory

Further Readings

Atkinson, R. C., & Shiffrin, R. M. (1971). The control of short-term memory. *Scientific American*, *221*, 82–90.

Baddely, A. (1990). *Human memory: Theory and practice.* Boston: Allyn & Bacon.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63,* 81–97.

Terry, W. S. (2003). *Learning and memory: Basic principles, processes, and procedures.* Boston: Allyn & Bacon.

SINGLE VERSUS COED GENDER EDUCATION

Single-gender education refers to the delivery of instruction to members of the same sex in isolation from members of the opposite sex, whereas *coed gender education* refers to instructional delivery to males and females simultaneously. Although gender grouping is done at both schoolwide and classroom-specific levels, this entry focuses on the former.

The issue of single versus coed gender education has practical implications for schooling, yet most stakeholders' rationale for making a choice between options has more often than not been an issue of politics and economics rather than pedagogy. In the modern era, but prior to the 20th century, formal education at all levels was decidedly single gender—worldwide. Coeducation has since become common practice and is more often the instructional option of choice, especially among public educational institutions. In the early 1900s, public school coeducation across the globe was advanced through the promise of political equality for democratically educated citizenries. Less idealistic but no less evident, coeducation in the West was used as a vehicle for cultural assimilation. It was a response to the waves of European and Asian immigration that threatened to make over America's colonial heritage. Furthermore, the Depression made compulsory education and coeducation emblematic of good business. After all, the logic went, coeducation required fewer schools, teachers, and support staff.

With the national turn toward coeducation, single-gender education became the province of private, often religious schools, especially in the United States. Catering to members of elite communities, private and religious schools monopolized single-gender education. The nation's elite private and religious single-gender schools filled their classrooms by suggesting the firestorm of adolescent courtship detracted from the essential academic tasks of schooling until declining economic circumstance in the 1970s and 1980s forced them to join the march to coeducation. Toward the end of the century, however, legislation and policy in several countries, including the United States, Great Britain, Australia, and New Zealand, reflected a resurgence of interest in private and public single-gender education.

Current Trends and Limitations

To date, empirical investigation of the issues surrounding and outcomes of single versus coed gender education has been sparse in comparison to topics such as gender differences in social behaviors and academic attainment. What has been conducted has been void of experimental designs or includes anecdotal information from single schools or districts. In this sense, consideration of gender or sex as a grouping variable is regarded as another on a list of group manipulations no more potentially relevant to the average student than grade level or socioeconomic status. Among research that is relevant, however, results are mixed. Support for either position is couched in one of two veins. The traditional argument for coeducation is that separation is inherently unequal. Moreover, male and female social interaction during school years helps to secure a foundation for the future of democratic society. A different perspective borrows ideas from sociobiology and the cognitive sciences and is committed to single-gender schooling, noting that, on average, males and females learn differently and socialize differently. Gender segregation, it is argued, is essential to optimize the academic potential of females and males.

In spite of—or maybe because of—the paucity of research, advocates of both sides of the debate cite research supporting their broadly staked positions. Proponents of single-gender education tout studies with findings that suggest single-gender education enhances gender self-esteem, is beneficial for male discipline, and is well received by teachers. Most importantly, at least for advocates of current accountability, single-gender education is suggested to enhance attitude and achievement in core subjects, including math, sciences, and language arts, as well as physical, health, and technology education. Coed gender education advocates counter by calling attention to research that suggests coeducational environments are essential to ensure equality of educational opportunity, diminish sexism in the classroom, end gender discrimination, and circumvent

the glass ceiling that once excluded women from all of the best educational institutions and profitable professions.

The inability to resolve many of the issues in the debate over single-gender versus coeducation instructional environments suggests that research methods need reexamination. Reliance on sociological instead of experimental methods has confined much of the conversation of the topic to personal perspective and observation. Such methods also result in an overreliance on the investigation of idealized behavioral outcomes (i.e., grade point average, standardized test scores, discipline referrals, attendance). Inasmuch as these outcomes are significantly downstream from the intervention of single-gender or coeducation, the inferential value of any concerned studies is limited. Moreover, blanket statements of the effectiveness of one or the other suggests that a culture of single-gender or coeducation is a bounded system made up of philosophies and practices that can be transmitted en corpus to students. Here, single-gender or coeducation is reified in favor of identifying the advantages and limitations of each instructional type for individual students.

Future Directions

Findings from the cognitive sciences reveal that physical organization of male and female phenotypes are two very similar yet profoundly different physical and mental code sets through which each individual recognizes his or her world and physical relationship to it. This encoding presents not just a physical differentiation but also a differentiation of the cognitive, affective, and conative structures for the mind. To date, for example, differences between males and females have been found in a host of informational processing and behavioral measures, including advantages for males, on average, over females on a number of spatial skills such as mental rotation of figures, maze solving, estimation of speed for an object in motion and horizontal perception. Female advantages, on average, over males have been found on depth perception, fine motor dexterity, and location memory tasks. If gender of classmates is related to academic or social outcomes of students, then intraindividual differences in brain function should be evidenced when in the presence of males versus females. In this sense, using the research methods developed by researchers in the cognitive sciences for experimental

investigation of intraindividual response to varying gender of classmates may help to recast consideration of single and coed gender education as something more than a pair of monoliths. As these findings are examined, a more precise, and ultimately practical, discussion is likely to arise from the single versus coed gender education debate.

Sean Alan Forbes and James S. Kaminsky

See also Cultural Diversity; Gender Bias; Gender Differences; Gender Identity

Further Readings

Mael, F., Alonso, A., Gibson, D., Rogers, K., & Smith, M. (2005). *Single-sex versus coeducational schooling: A systematic review.* Washington, DC: U.S. Department of Education.

Pinker, S. (2002). *The blank slate: The modern denial of human nature.* New York: Viking.

Salomone, R. C. (2003). *Same, different, equal: Rethinking single-sex schooling.* New Haven, CT: Yale University Press.

Salomone, R. C. (2006). Single-sex programs: Resolving the research conundrum. *Teachers College Record*, *108*(4), 778–802.

Tyack, D. B., & Hansot, E. (1992). *Learning together: A history of coeducation in American public schools.* New York: Russell Sage Foundation.

SOCIAL CLASS AND CLASSISM

Social class and classism is an important and relevant topic for educational psychologists because extensive research shows that children and adolescents are affected by their social class and experiences with classism. Although the research sometimes has problems conceptualizing and measuring social class in consistent and meaningful ways, nevertheless, the body of evidence suggests that educational psychologists need to consider the student's social class context. For educational psychologists, the "tip of the iceberg" is the student in the classroom. Beneath the manifest student is the latent weight of the student's social class. He or she brings into the classroom the affects of family, peer, and environmental social class. Educational psychologists need to consider how poverty, for instance, is related to physical growth, cognitive stimulation, and intellectual development. Depending on what social

class criterion is being measured, different results will arise. Therefore, approaching social class assessment with a coherent theory or research question will be an important first step.

In this review, several key considerations and issues are discussed. First, social class, classism, and socioeconomic status are defined as general psychological concepts. Second, the ways in which social class and inequality have an impact on students before entering the classroom environment is discussed. Finally, specific issues such as the achievement gap, teacher perceptions and expectations, and teacher quality are presented.

Defining Social Class and Socioeconomic Status

Social class may be regarded as a cultural construct and part of a psychologist's multicultural competency. But insofar as social class and classism are ubiquitous in an individual's life span, psychologists need to understand it as more than just a cultural variable. For many psychologists, the main difficulty of using social class is definitional. A few questions may be: What is the difference between social class and socioeconomic status? How is it measured? Is social class a position, or is it malleable?

To begin, several researchers have suggested that there is no real conceptual difference between social class and socioeconomic status (SES). The variations in research are a result of different conceptualizations by researchers as well as the use of differing variables to measure their operationalization of social class or SES. In a review of three counseling and psychology journals across 20 years (1981–2000), Liu and his colleagues found authors used more than 400 different terms to conceptualize social class or SES (e.g., inequality, poverty). Similar results have also been found in educational psychology journals where social class is seldom measured or meaningfully analyzed. Additionally, researchers have varied in defining social class or SES as access to money, resources, mobility, or power and have also speculated that social class and SES are static positions related to inequality or poverty. For the purpose of this entry, social class is used throughout this entry because it is nominally easier to link with classism. Also, social class reflects the stratification that is both the cause and consequence of inequality. Additionally, people

are likely to interact with others based on perceived social class differences.

Social class is also co-constructed with classism much like race and racism are interdependent constructs. That is, the operationalization of race cannot exist without the explicit or implicit function of racism to create rigid categories based on phenotypes or other manifest physical features. Similarly, social class cannot exist without the function of classism to create stratification and inequality. Individuals would not categorize other individuals into varying social class groups if there were no privilege or gain associated with a hierarchy of groups. In this case, classism functions to unequally distribute privilege and gain within and across various social class groups. Therefore, it is important to discuss social class and classism together to reinforce the relationship between social class position and the inequality that creates it.

Related to terminology is the measurement of social class. Typically, psychologists have adopted a sociological framework to understand social class. The problem is the macro level of analysis and understanding in sociology (i.e., understanding how groups, communities, and societies function) versus the individual level for psychologists. Consequently, rather than examining interpersonal or psychological variables associated with social class, psychologists have used income, education, and occupation, either individually or variously aggregated, to categorize individuals into discreet social class groups (e.g., $100,000 in income, a college education, and a white-collar occupation = middle class). The problem with this categorization approach is that the variables are not highly intercorrelated, and there is no evidence to suggest that a particular income, education level, or occupation will effect or produce a specific social class. Furthermore, psychologists need to understand that income, education, and occupation are associated with different psychological and educational outcomes. For instance, income is notoriously difficult to measure because people do not give accurate answers. Income is a good predictor of future career success, but more robustly for the poor than for the rich. Another problem in using these variables is the assumption that everyone within a specific social class group perceives the world similarly. But evidence suggests that, even within the same economic context, people vary in how they conceptualize their environment. For instance, psychologists with the same degree, education, occupational experience, and income will likely

see the world differently if one lives in Boston, Massachusetts, versus Iowa City, Iowa. Both psychologists may regard themselves as middle class, but being a doctoral-level psychologist may have a different status within a largely rural community versus a large urban community.

For educational psychologists, by using adult indices such as income, education level, and occupation, one may conclude that social class and classism are adult experiences. Yet research suggests that children and adolescents are aware of and able to make social class discriminations such as what constitutes lower and middle class. In addition, children and adolescents are sensitive to social class prejudice and classism. For instance, in one study of African American boys in a private school, the researcher found that these boys struggled to negotiate Black and White cultural norms in school so that they could succeed in school and not be perceived as a "sellout" by other African American peers. Similarly, children and adolescents are sensitive to advertisements and are oriented to buying material objects consonant with their perceived social class standing. Therefore, psychologists need to be aware of methodologies used to assess social class at different age levels. Unlike race, which is a constant feature, unlikely to change with age, social class is much more likely to change over time. Ultimately, there is no one best approach to measuring social class, but it does depend on the theoretical operationalization and research questions.

Burgeoning research by Nancy Adler and Michael Marmot and colleagues also suggests the importance of subjective appraisals of social class. That is, rather than asking for income, education, and occupation, ask an individual to place him- or herself on a 10-rung social class ladder where the bottom rung is the lowest position in society and the highest rung is the top of society. Evidence in health psychology research shows that this subjective method may be a robust and good predictor of self-rated health. Liu and colleagues have also developed a social class worldview model (SCWM) predicated on a phenomenological approach to understanding a person's social class and experiences with classism. The SCWM uses the worldview framework to interpret social class experiences and uses social class as a descriptor of a psychological process. Psychologists should continue to use social class as an adjective (descriptor) of a meaningful psychological process such as identity, acculturation, stress, and strain.

Additionally, Liu and colleagues describes a network of different classisms such as upward (prejudice against those perceived to be in a higher social class), lateral (pressure to "keep up with the Joneses"), and internalized (feelings of anxiety and apprehension related to losing one's social class standing). Although subjective and phenomenological approaches to measuring social class are important developments, there are still objective consequences to low income, poverty, and social inequality. Thus, experiential methods may illuminate the interpersonal experiences, but psychologists still need to contend with the contextual issues related to social class.

The Impact of Social Class and Inequality

A network of factors converge to create inequality, poverty, and classism. Psychologists need to consider sociopolitical (e.g., the unequal distribution of power); sociohistorical (e.g., biased and inaccurate histories of peoples); and sociostructural (e.g., legal, educational, and economic systems) forces that create inequalities. The result of these inequalities is a gradient, which generally suggests that children and families at the lower social classes have poor physical and mental health and low school achievement, whereas those in the higher social classes have just the opposite. The most pervasive manifestation of social class in educational settings is the inequality resulting from inadequate resources. This lack of resources starts before the child enters school and is only exacerbated in poor school settings.

Closely related to educational disparity is the impact of race and racism. Social class, classism, and racism are often conflated in the research, but this conflation likely reflects the interdependency and relatedness of racism and classism rather than methodological limitations. Extensive research evidence shows that racial and ethnic children and adolescents (e.g., African, Asian, Latino, and Native American) are more likely to live in households with disadvantage, low incomes, high incidents of unemployment, violent and environmentally toxic neighborhoods, and poor health-related resources such as supermarkets and hospitals. These poor family and neighborhood environments have also been correlated with physical health problems such as obesity, diabetes, and poor cognitive and intellectual development. The National Institute of Child Health and Human Development Early Child Care Research Network found that in school settings, the aggregate

result of health, social, and environmental deprivation and stressors may be dropping out of school early and poor socioemotional development and competence, as well as other achievement gaps.

Achievement Gap

In educational psychology, researchers have identified disparities in academic achievement between students from lower and higher social class backgrounds. For instance, in one 12-year longitudinal study examining several risk factors (e.g., low education, low occupational status) and their relationship with students' future academic achievement, researchers found that the students whose families had low educational backgrounds and low occupational status were more likely to perform worse on intelligence tests, have lower grade point averages, and a greater number of absences than students who were not as disadvantaged. Researchers have also found that lower social class is related to higher dropout rates and enrollment in special education classes. In another study, the National Assessment of Education Progress found that students from lower social class backgrounds are also less likely to master academic skills. Differences in achievement have also been found when individuals increase their social class status. In one study, children from single-mother households perform better on cognitive assessments when their families are able to remain off welfare. It is important when discussing the problems related to lower social class backgrounds to remember the array of contextual problems affecting performance and not focus on dispositional attributions (i.e., blame the child). Specifically, contributing to poor performance may be early exposure to lead and other toxins, poor schooling, reduced access to affordable healthy food, neighborhood violence, and dilapidated playgrounds, to name a few.

Furthermore, researchers caution that multiple variables may be influencing the relationship between social class and academic achievement. Factors typically associated with lower social class, such as racial and ethnic minority status, and greater levels of stress can influence the relationship between social class and student academic achievement. Racial and ethnic minority students must contend with additional pressures such as discrimination. Also, students who have limited access to health care may be likely to miss school and less likely to receive proper care for illnesses or injuries that modify cognitive abilities.

Finally, students who are under constant stress may not prioritize schoolwork as highly as students who experience less stress.

Teachers' Perceptions and Expectations

Research suggests that teachers' perceptions and academic expectations of students may alter as a function of social class. Teachers may perceive lower social class students' cognitive abilities more negatively than higher social class students' cognitive abilities. These negative perceptions may translate into lower student expectations because teachers expect less from students whom they perceive as less than capable. Students from lower social class backgrounds may also experience lower teacher expectations than students from higher social class backgrounds. As a result, teachers may differentially treat lower social class and higher social class students in the classroom. This differential treatment suggests that lower social class students may face additional barriers than do their higher social class counterparts. For example, teachers with low expectations of students are more likely to demonstrate negative biases toward those students, ask them fewer questions during class, provide less positive reinforcement, and offer less support in comparison to students whose teachers expect more from them. The focus on teachers' perceptions and expectations has become an important topic because teachers with low expectations of students are more likely to have lower-achieving students. Several factors have been identified that may influence the relationship between social class and teachers' perceptions and expectations.

Moreover, researchers have argued that teachers' negative perceptions of lower social class students may lead to the development of self-fulfilling prophecies. Lower social class students may internalize the lower expectations of teachers and achieve academically in a manner that is consistent with those teachers' expectations. Recent studies have shown that self-fulfilling prophecies may adversely affect student achievement, but the impact of self-fulfilling prophecies on student achievement may be minimal.

Teachers' self-efficacy for working with lower social class students may also be negatively related to their perception of students. Teachers with low self-efficacy may be more likely to identify special education as more suitable for lower social class students

with academic problems than teachers with high self-efficacy. Teachers with high self-efficacy may have greater confidence in their ability to work with students with academic problems irrespective of their social class background.

Teaching Quality

Researchers have argued that social class is related to teacher quality. School districts comprising primarily lower social class neighborhoods are more likely to have less qualified or uncertified teachers. Schools in lower social class school districts may have fewer resources to recruit and maintain highly qualified teachers. Teachers in lower social class school districts may also be limited by a lack of instructional materials such as computers or lab equipment. In addition, the Education Trust found that teachers who scored in the lower range of national teacher qualifying examinations are likely to teach in lower social class schools. The incessant need for teachers in lower social class schools may require those schools to be less selective in their recruitment of teachers. Finally, teachers in less affluent school districts are more likely to teach subjects unrelated to their field. For example, teachers who hold degrees in English education may be required to also teach history.

More recently, the introduction of No Child Left Behind (NCLB) may create greater discrepancies between higher and lower social class schools. Highly qualified teachers may be more inclined to leave underperforming, lower social class schools because of NCLB regulations that penalize underperforming schools. As a result, students with less qualified teachers are more likely to experience lower levels of academic achievement than students with highly qualified teachers.

William Ming Liu and Jovan Hernandez

See also Cultural Deficit Model; Cultural Diversity; Ethnicity and Race; Poverty

Further Readings

Adler, N. E., Boyce, T., Chesney, M., Cohen, S., Folkman, S., Kahn, R., et al. (1994). Socioeconomic status and health: The challenge of the gradient. *American Psychologist, 49,* 15–24.

Arnold, D. H., & Doctoroff, G. L. (2000). The early education of socioeconomically disadvantaged children. *Annual Review of Psychology, 54,* 517–545.

Duncan, G. J., Yeung, W. J., Brooks-Gunn, J., & Smith, J. R. (1998). How much does poverty affect the life chances of children? *American Sociological Review, 63,* 406–423.

Evans, G. W. (2004). The environment of childhood poverty. *American Psychologist, 59,* 77–92.

Liu, W. M., Ali, S. R., Soleck, G., Hopps, J., Dunston, K., & Pickett, T., Jr. (2004). Using social class in counseling psychology research. *Journal of Counseling Psychology, 51,* 3–18.

Liu, W. M., Hernandez, J., Mahmood, A., & Stinson, R. (2006). The link between poverty, classism, and racism in mental health. In D. W. Sue & M. G. Constantine (Eds.), *Racism as a barrier to cultural competence in mental health and educational settings* (pp. 65–86). New York: Wiley.

McLoyd, V. C. (1998). Socioeconomic disadvantage and child development. *American Psychologist, 53,* 185–204.

Murdock, T. B. (2000). Incorporating economic context into educational psychology: Methodological and conceptual challenges. *Educational Psychologist, 35,* 113–124.

National Assessment of Education Progress. (2000). *The nation's report card: Fourth grade reading 2000.* Washington, DC: National Center for Education Statistics.

National Institute of Child Health and Human Development Early Child Care Research Network. (2005). Duration and developmental timing of poverty and children's cognitive and social development from birth through third grade. *Child Development, 76,* 795–810.

SOCIAL DEVELOPMENT

Children's *social development* refers to the ways in which children grow in terms of their social skills, awareness of others, cooperative behaviors, and ways of approaching and interacting with others. Children's social development has significant implications for later functioning in numerous areas, including emotional development, educational and employment success, and overall adjustment. Poor social development can place children at risk for poor relationships with their peers, academic problems, criminal activity, and mental health and adjustment problems. Thus, successful development socially is integrally important to children's overall well-being.

From the moment they are born, children enter and have to begin learning how to navigate a social world. A graphic model of this is given in Figure 1. Early in life, children's primary social context for developing is their immediate family. Later, as they develop, children become increasingly more integrated into other divergent social contexts, including their peer groups,
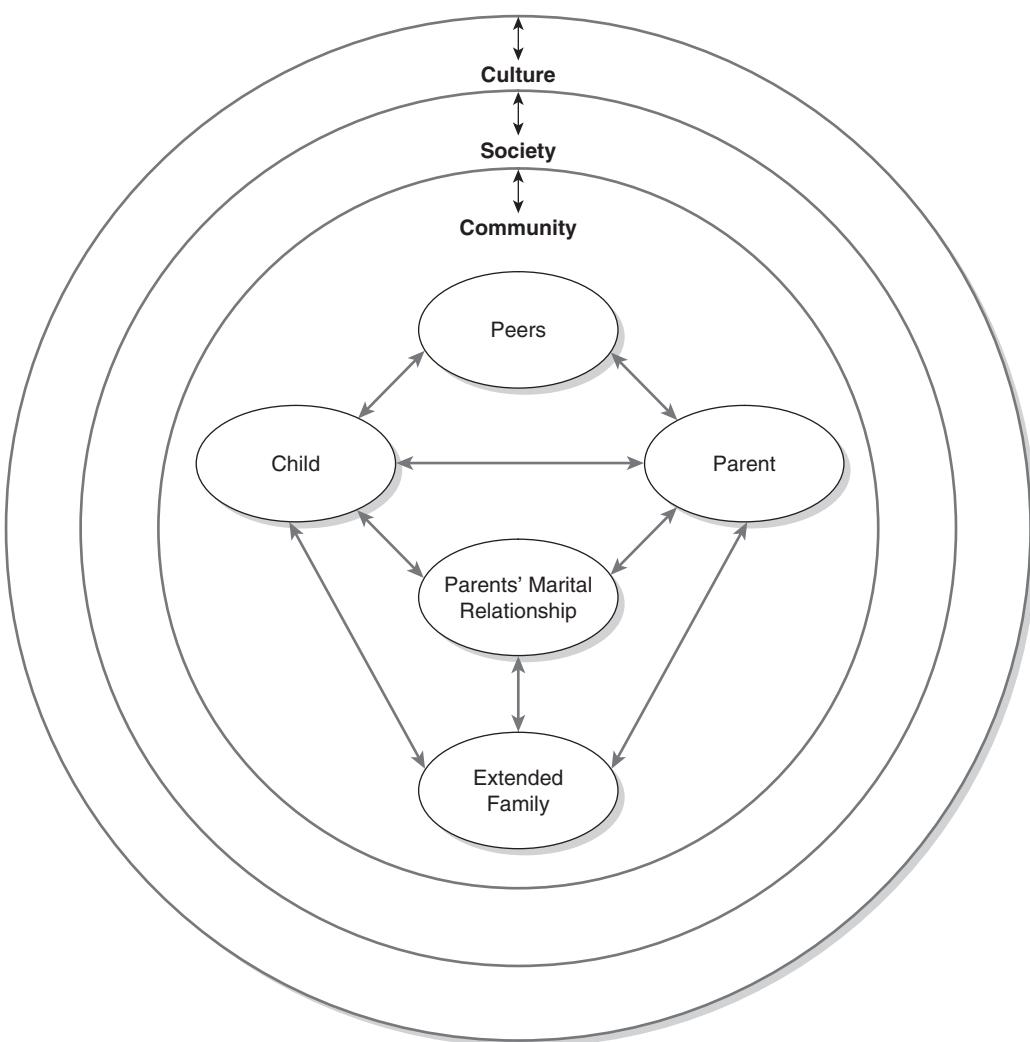
Figure 1 Model for Social Development

Source: Adapted from Lerner, R. M. (1979). A dynamic interactional concept of individual and social relationship development. In R. L. Burgess & T. L. Huston (Eds.), *Social exchange in developing relationships* (pp. 271–305). New York: Academic Press.

school networks, professional networks once they enter the job market, and eventually their own romantic relationships. Throughout their development over time, they are also affected by more distal influences such as the community, society, and culture in which they are living.

Family Context

Parent–Child Relationship

Early in children's lives, their primary social interactions are with their primary caregivers, typically their mothers and fathers (although this can vary within and between cultures), and these early social interactions set the basis for all interactions later on. Infants learn important lessons about what to expect from others and how to meet their needs through their repeated interactions with their parents over time. The focus of early social interactions for infants surrounds having their basic physical and emotional needs met. All infants vary in their emotionality and their need to be soothed, as well as their abilities to soothe themselves and be soothed by others. Differences in parenting styles, such as insensitivity, unresponsiveness, and inconsistency may lead children to be more

emotional and have greater difficulty in regulating their emotions, because they do not know what to expect and are not having their own needs met appropriately. More emotional children may elicit more negative responses from their parents, and thus a negative reciprocal cycle is established.

John Bowlby and Mary Ainsworth contend that infants' patterns of interactions with their parents develop into patterns of attachment style. *Attachment* refers to the enduring emotional bond that develops between people. Parents who are sensitive, responsive, and consistent in meeting their infants' needs over time are more likely to have infants who develop a secure emotional attachment. When parents are unable to consistently meet their infants' needs, then an insecure emotional attachment is more likely to develop. Children who have a secure attachment feel confident that their parents are there for them when they need them and are able to successfully reach out to their parents in times of physical, social, emotional, or other needs and be comforted. On the other hand, children who are not securely attached may not feel safe or comfortable seeking help from their parents because they have not successfully had their needs met in the past. Instead, they may try to hide their emotions or needs. When insecurely attached children do attempt to seek out their parents for help, they may either outright reject their parents' attempts at consolation or simply may not feel comforted by them.

Bowlby and Ainsworth argue that these early attachment patterns form the basis for all other relationships that children will develop. Children use their early attachment relationships with their parents as the interpretive lens through which to view, interpret, and respond to all other relationships. In other words, they develop internal representations or schemas about how relationships work based on their relationships with their parents. Levels of parent–child attachment security provide a global sense of emotional security that influences children in many broad domains. Children seek to recreate with peers, and later with intimate romantic partners, the relationship they had with a parent. Children's levels of security and trust from their attachment relationships determine how open and trusting or wary and apprehensive they will be in approaching a new person with whom to develop a relationship. Thus, children who are securely attached are more likely to also have other positive, successful relationships with other people, whereas children who are insecurely attached are more likely to demonstrate difficulties in interacting with others. The patterns of interaction styles between parents and infants around their emotional needs set the stage for children's future interactions with others.

As children grow older toward the preschool years, they begin to become more aware of social tasks and demands and are capable of initiating, maintaining, modulating, or ceasing physical acts, communication, and emotional expressions as required. Differences in social development may be attributable to differences in parenting. Children whose parents have been consistent in teaching them what is socially appropriate behavior versus inappropriate behavior may be more likely to conduct themselves in more socially desirable ways. Furthermore, children who have successfully had their emotional needs met early on are better at regulating their own emotions as they develop and thus have more success in modulating their outward expressions of emotions, such as their behavior and other communication, because they can better modulate their internal states.

Marital Relationship

The parents' marital relationship is another key factor contributing to children's social development. The quality of the marital relationship and the ways in which marital conflict is handled set the emotional climate for the family and the child. Marital conflict has both positive and negative elements, however, and depending on how it is handled, it can have either positive or negative effects on children's social development. Heightened levels of negative conflict can influence children directly and indirectly. Greater conflict between parents may lead to greater disturbance in parenting, which can lead, in turn, to less emotional security and the ensuing issues described above.

Increased conflict can also directly influence children, leaving them feeling vulnerable and emotionally insecure about the stability of their family. E. Mark Cummings and Patrick Davies proposed an emotional security hypothesis, which extends the ideas from attachment theory and suggests that children can develop a sense of emotional security not just to the parent–child relationship, but also to the marital relationship. From their hypothesis, children react to the meaning of conflict itself, rather than just the presence of conflict, and the more they are exposed to it, the more sensitized they become to it over time. Witnessing more negative, hostile, and threatening forms of

conflict is particularly detrimental to children's sense of security about the marital relationship, whereas more positive expressions, including humor, support, and affection, promote greater feelings of security.

As in attachment theory, children's sense of emotional security about the marriage can also influence children's relationships in other realms, and children who are more emotionally insecure about the marriage have more social difficulties with other children. Children also learn ways in which to handle conflict and other social interactions based on how their parents interact and handle their own conflict and relationship. When parents engage in more destructive and hostile forms of conflict, children are more likely to hold negative views of peers and to act similarly with peers. For example, when children witness acts of aggression between parents, it results in greater physical aggression, such as hitting and pushing, toward their peers.

Other Family Relationships

The quality of parent–child and marital relations has significant implications for broader family functioning. Positive parent–child and marital relations are related to more harmonious sibling relationships, whereas discordant marital relations are linked with more negative sibling interactions. Other family relationships beyond the parent–child and marital relationship can also contribute positively or negatively to children's social development. Some cultures have different definitions for what constitutes a family, and/or different traditions for living arrangements, all of which provide a larger family group with which children can interact on a regular basis. When there are negative parent–child or marital relationships, having strong, positive healthy relationships with other family members, such as grandparents and aunts and uncles, can help ameliorate negative consequences for children's social development by providing other positive role models and sources of support and emotional security. On the other hand, when these relationships are also problematic, it can further exacerbate social interaction difficulties for the child.

Sibling and cousin relationships provide some of the first opportunities to practice emerging social skills in similar aged, peer-like relationships. Many tribal societies group children into age-grade peer groups, with whom they spend the majority of their waking hours, much like the school system in industrialized societies.

Allowing children to work through their own disagreements with siblings is important for the development of appropriate social skills, and frequent intervention on the part of parents in sibling conflict can interfere with this process. Through these relationships, children learn more about appropriate social norms and deviance and begin to put into practice their knowledge of relationships learned from their parents as they develop specific social skills such as cooperative behavior, sharing, assertiveness, social conversations, empathy, problem solving, conflict resolution, and issues of acceptance and rejection. These are the skills on which children will need to draw to be successful in their interactions with peers outside of the family context.

Peer Groups

Peer Social Status and Friendships

In developing peer relationships, children bring with them the lessons they learned and the relationships they internalized from their family environment, which has a large impact on their acceptance versus rejection by peers and their development of friendships. Steven Asher and colleagues suggest that both acceptance and friendships are important for healthy adjustment and development. They define peer acceptance as the extent to which a child is liked or accepted by other members of the peer group. Well-accepted children are warmly and positively regarded by most of their peers, whereas poorly accepted children tend to be viewed negatively and disliked by their peers. Acceptance is thought to be a unilateral concept, which works in one direction. Friendship, on the other hand, is more bidirectional, and friends perceive and respond to each other as unique and irreplaceable. Acceptance is thought to be important to the development of healthy attitudes toward competition, conformity, and achievement, whereas friendship is key in the development of empathy and perspective taking, and it validates children in terms of their interests, positive self-views, and hopes. Children's acceptance and friendship relationships are significant predictors of their long-term well-being and success, and children who are not successful in these areas are more likely to have adjustment problems (e.g., anxiety and depression symptoms), drop out of school, describe feelings of loneliness, display aggressiveness and/or submissiveness, and experience social ostracism and isolation. Estimates of the percentage of

children experiencing serious difficulties with their peers are around 10%.

John Coie put forth a theory of child rejection. Assumptions of this theory are as follows:

1. Social behavior is primarily responsible for rejection by peers.

2. The difficulties of rejected peers result from the way they interpret specific social situations, the way they react affectively, and their acquired strategies for dealing with the situations.

3. These factors emerge out of their socialization history and are largely shaped by the history of the child's interactions with parent figures, siblings, and nonsibling acquaintances.

The theory suggests that there are four stages of social rejection:

1. the *Precursor* stage, in which children come equipped with specific competencies and deficits relating to their eventual peer status;

2. the *Emergent* stage, in which their interactions with a significant peer group result in the child being rejected;

3. the *Maintenance* stage, in which rejection by a group becomes a stable and enduring reality;

4. the *Consequence* stage, in which other aspects of the child's adjustment deteriorates to the point of identifiable psychological disorders.

In the Emergent stage, children who are more aggressive and disruptive are more likely to be rejected, as are children who are more withdrawn and solitary. Children who become socially successful and popular take the time to figure out what is happening in a new group situation and attempt to match their behavior to that of the group, whereas children who are less popular are more likely to disrupt the group in part because they are not aware of what is happening within the group.

In the Maintenance stage, the rejecting group members begin to change in their behavior toward the rejected child, and in turn, the rejected child begins to change in his or her behavior toward peers, his or her feelings about him- or herself, and in the thoughts and expectations for the self and others. Being rejected begins to be a part of the child's identity, and the child and others begin to respond now with new expectations. Once rejected, rejected children are more likely to be

on the outside of well-established, coherent playgroups and even have trouble getting along with other rejected children. Rejected children describe themselves as being more lonely and having lower self-esteem, and when rejected children are also aggressive, they have concerns about the need to prevail against others, whereas nonaggressive children develop more concerns about being scorned and attacked by others.

In the Consequences stage, children who were rejected in middle childhood and preadolescence are at greatest risk for later disorder because childhood is the time for the preparation for intimate relationships later on. Rejected children have more academic difficulties, are truant more often, have more discipline problems, drop out of school at higher rates, and are left with fewer social skills to develop appropriate relationships later on in development.

Children who are lacking in the necessary social skills associated with acceptance and friendship have long been the focus of psychological assistance. Asher and colleagues review numerous behaviors that have been the target of interventions aiming to improve children's social skills, ranging from broad general skills to more specific molecular skills: social problem solving, nonverbal communication, negotiation, entrance to groups, ability to cope with anger, helping behaviors, establishment of rapport, conversational skills, listening skills, self-control, ignoring of bullying and rejection, expression of requests and rights, ability to give and receive compliments, sportsmanship, and so on. This multitude of skills has taken place in various modalities, such as role-playing, modeling of successful peers, videotaped interactions that are reviewed and discussed, and actual interactions with peers. In order to have a successful intervention, it is imperative that children are able to practice the learned skills in actual interactions with their peers. The majority of these interventions are successful in increasing children's peer acceptance. It is still not clear, though, why the behaviors learned in these interventions translate into greater success in the friendship realm.

Asher and colleagues suggest that in order to develop strong, lasting friendships, children need social skills that are adaptable across various settings and need to interact with peers in multiple settings to create more multifaceted and more invested relationships. Additionally, they need the skills and personality necessary to be perceived as fun, resourceful, and enjoyable companions; must have reciprocity in their

relationships; must be able to engage appropriately in self-disclosure; must be able to express caring, concern, admiration, and affection in appropriate ways; must be helpful and reliable when their friends need them; must be able to manage minor disagreements successfully while preventing major conflicts from occurring; must be able to forgive; and must be able to handle outside influences (such as school and other children) on their friendship relationship.

Bullies and Victims

Although peer rejection places children at significant risk for the development of a wide range of problems, being the victim of a bully creates an even greater risk. Schwartz and colleagues conducted a study comparing two different types of victims of bullies: aggressive victims and passive victims. They also compared them to aggressive but nonvictimized children. They speculated that there would be different socialization histories among the different groups. Studying the three groups over time, they found that aggressive victims had preschool experiences with harsh, disorganized, abusive home environments; maternal hostility; restrictive and overly punitive parents; and higher interparental conflict. Furthermore, 38% of the aggressive victims of bullies were physically harmed by their parents and/or other adults. In comparison, nonvictimized but still aggressive boys were also exposed to high amounts of adult conflict and aggression, but had not been abused or exposed to harsh treatment by their parents. Finally, the passive boys showed no differences in socialization experiences compared to normal boys and also experienced less overall exposure to aggressive socialization factors. Their results thus support the importance of early socialization histories within the family in terms of predicting victimization by bullies later in life.

Just as there are various types of victimized kids, there are also various types of popular children. Surprisingly, it is not only prosocial (model) children who are popular; a certain subset of more antisocial and tough boys is also highly popular. For example, an intriguing study by Philip Rodkin and colleagues found that prosocial model boys were perceived by their peers as cool, athletic, leaders, cooperative, studious, outgoing, and nonaggressive, and were perceived by themselves as nonaggressive and academically competent. Overall, the model boys were highly popular. Tough boys, in contrast, viewed themselves

as popular, aggressive, and physically competent, and they were more likely to be of a minority status, particularly African American. The tough boys were among the most popular and socially connected children in the school by their perceptions as well as by the other students' and teachers' perceptions. Thus, popular boys are a heterogeneous group, including both model and antisocial members. The authors argue that the peer cultures of the two different popular groups differed in important ways. Some minority status groups value behaviors such as academic disengagement and disobedience of school rules that are in opposition to dominant societal preferences. Alternatively, the aggression may be functional for those African Americans who are socialized in low-income and higher-risk communities, in that it is desirable and perhaps keeps them safe in such situations. The visible European American presence and values may make it difficult to be academically successful, popular, and prosocial without "acting White." Thus, devaluing prosocial behavior may be a viable means for them of obtaining social status.

Broad Distal Influences

Distal, Not Distant

The effects of community, society, and culture historically have been discussed as "distal" because people do not typically have personal interactions with their culture the same way they do with their peers, and certainly not on a daily, moment-to-moment basis like they do with their families. Most people are not even cognizant of their culture, even in this self-aware global age. It is only when one is removed from one's culture that one becomes aware of it, or, as in the case of the indigenous people of most industrialized nations, when one's culture is surrounded (they might say "besieged") by another.

The effects of community, society, and culture may well be distal, but they are not distant. They surround everybody every day and influence them in much the same way that fish are influenced by water. Social development takes place within a community, which exists as part of a society. What constitutes a community or a society, among many other things, is defined by culture. Although different societies may have different sets of rules to which one must acculturate, every aspect that makes up the very definition of social development—social skills, awareness

of others, cooperative behaviors, and ways of approaching and interacting with others—is defined by culture and therefore can vary between cultures. Indeed, anything that has meaning that is shared among people is a result of culture: Meaning is defined by culture.

Culture

The concept of culture arose in the age of discovery as Europeans expanded across the globe and discovered that not everyone looked and acted like them. The original debate was between nature and culture, or "Are those brown people so different from us because God made them that way, or because they were raised in that heathen society?" The concept of culture has been widely accepted into the Western scientific worldview, but there is no standard definition of culture. For purposes of this entry, *culture* mostly refers to the set of values and beliefs that a group of people shares and with which the people make sense of their world.

Society

In most Western cultures, there is a strong identification between nationality and society, but this is a concept that is defined by culture and is by no means definitional or absolute. This concept is so ingrained in Western society that it is seen by most members as a natural part of the world, rather than something imposed upon it by their worldview. A list of examples such as this—things that seem "natural" but are, in fact, defined by culture—could extend nearly indefinitely. Indeed, anything that is a value judgment is so defined by culture.

In other cultures, a society may consist of many villages or perhaps only one. A society could even be the totality of a culture, but this seems unlikely in the world today. The Iroquois Confederacy is a good example of an indigenous society that was much like our own, in that it was made up of many villages and even members of several different cultures. The Kurdish people today are a good example of a society that shares a culture but not a single nation, although the latter is mostly an artifact of the colonial era and soon may no longer be true.

Western society is multicultural, subsuming the members of many different cultures and subcultures. Members of different cultures become part of Western society mainly by immigration, and acculturation to Western culture varies across groups and individuals.

The "melting pot" ideal of early 20th-century America has mostly been abandoned, and the more realistic "stew pot" notion—which allows ingredients to flavor the stew while still remaining recognizably distinct—has become increasingly recognized.

Subcultures are groups within the society that share a worldview and belief system that is different from the main culture. It is quite possible that some subcultures may not even consider themselves part of the larger society. Rastafarians and their reference to the Western societies they live within as "Babylon" is a good example of this phenomenon. How we can say that these subcultures are, in fact, part of the larger society is beyond the scope of this entry, but the short answer is mostly geographic—they are surrounded by and/or live within the main culture—and they also have frequent interactions with their non-subculture neighbors—but this topic is open to interpretation.

Community

Community is usually defined as a fairly tightly knit group that has a sense of itself as a group, whereas a *society* is a much larger, loosely knit group that may consist of many communities and may also consist of subcultures that do not consider themselves part of the larger society. The model in Figure 1 refers to the social world of one individual, in which all of the overlapping peer groups discussed above comprise the community of each individual. All of the overlapping groups of all of the individuals in all of the peer groups comprise the actual physical community (but this would be an unreadable jumble of overlapping circles if they were represented in this format), and all of these communities together form the greater society.

Native American reservations are a rather unique subset of communities that defies most of the tenets set out here. They are physically within American society, but also separate from it—both physically and culturally. People who traverse the borders of Indian reservations come closest to actual interaction with culture as a real entity and not as an abstract concept. How one is defined as a person changes when one crosses that border. A man who is valued by his culture and revered as an omen of good fortune by the elders of the community can suddenly become an object of ridicule and scorn as soon as he enters White society. For most Native Americans, the transition may not be quite so dramatic as it is for this one individual, but would at the very least involve

awareness of the differing ideals and norms between the two worlds they inhabit.

Such negotiations between and across cultures provide unique challenges for individuals who must make them. Such people have more than twice the challenge: They must not only learn two sets of norms, but they must be able to successfully make the transitions between them. Healthy development in such a difficult environment requires support of a unique quality. There is a reason that Native Americans have the highest suicide rates in Western society, as well as a variety of other mental health issues, not the least of which is substance abuse. Many of the problems endemic in Native American communities today are the same as the problems that arise in any context of poverty and rural remoteness, but there is the added layer of cultural differences and the transitions mentioned to make life more difficult.

Native Americans are unique within this discussion in no small part due to their history: They are not a subculture that has grown organically within the matrix culture, nor are they a piece of the stew thrown in by immigration; rather, they are remnants of the multiplicity of cultures that existed on this continent before an immigrant society moved in and pushed them to the fringes. They are similar to immigrant communities that choose to isolate themselves into enclaves of resistance, with the exception that their isolation was not voluntary, and the option of acculturation was not historically offered to them. It has seemed in the recent past that a pan-Indian identity may be rising, wherein the multiplicity of cultures may band together and gain strength as an identity unified by what it is not, but it remains to be seen if the differences of the past can be left behind.

Although Native Americans provide a unique example of particularly complex and often difficult cultural circumstances to consider in social development, the struggles arising out of their minority status, poverty, and cultural clashes are common across many different cultural groups. These broader issues, together with the more proximal individual and family factors, interact bidirectionally and are woven together over time in intricate ways, creating the fabric of children's social development.

Tina D. Du Rocher Schudlich and Joel D. Schudlich

See also Bullying; Culture; Emotional Development; Family Influences; Friendship; Parenting; Peer Influences

Further Readings

Anderson, B. (1991). *Imagined communities: Reflections on the origin and spread of nationalism* (Rev. ed.). New York: Verso.

Asher, S. R., & Coie, J. D. (1990). *Peer rejection in childhood.* New York: Cambridge University Press.

Asher, S. A., Parker, J. G., & Walker, D. L. (1996). Distinguishing friendship from acceptance: Implications for intervention and assessment. In W. F. Bukowski, A. F. Newcomb, & W. W. Hartup (Eds.), *The company they keep: Friendship in childhood and adolescence.* New York: Cambridge University Press.

Booth, A., & Crouter, A. C. (2001). *Does it take a village? Community effects on children, adolescents and families.* Mahwah, NJ: Lawrence Erlbaum.

Bowlby, J. (1969) *Attachment and loss: Vol. I. Attachment.* New York: Basic Books.

Kopp, C. B. (1989). Regulation of distress and negative emotions: A developmental view. *Developmental Psychology, 25,* 343–354.

Sutton, J., Smith, P. K., & Swettenham, J. (1999). Bullying and "theory of mind": A critique of the "social skills deficit" view of anti-social behaviour. *Social Development, 8,* 117–127.

SOCIAL LEARNING THEORY

Social learning theory is considered to be one of the most influential theories in the fields of education and psychology. Social learning theory identifies learning as the primary factor in a theory of human functioning and personality development that is based on cognitive, social-interactive, self-regulatory, and self-reflective capabilities and processes. In addition to the impact of social interactions, social learning theorists believe that the media have a great influence on human learning and development.

Historical Background

Albert Bandura and Richard Walters's seminal 1959 book *Adolescent Aggression* set in motion the early foundation for the development of Bandura's ideas on social learning processes. Bandura and Walters initially sought to integrate principles from psychoanalytic and behavioral learning theories. Results from these early studies provided Bandura and his colleagues the data to empirically refute the psychoanalytic explanation of aggression, which was grounded in Freudian conceptualizations of identification with the aggressor and

catharsis or anxiety reduction need. At the time, Bandura and Walters's work was generally congruent with Robert Sears's social learning theory. Similar to Sears, Bandura integrated psychoanalytic and stimulus-response assumptions in a theory of social learning; however, Sears emphasized more psychoanalytic principles such as drive theory.

Beginning in the 1950s, researchers began to not only challenge psychoanalytic theories, they also began to question some of the principles of behaviorism and traditional learning theories because of their inability to explain some aspects of human development such as language development or the acquisition of novel responses. In particular, Bandura argued against theories that relied on trial-and-error learning because he believed that this approach was much too tedious of a learning process to explain how people acquire new knowledge and skills. Contrary to behaviorist perspectives at the time, social learning theorists emphasized the social contexts of the learning situation. Early social learning theorists began to provide counterevidence to the behaviorist learning principle that children could not change their behavior without first being reinforced for approximations to the new behavior. Bandura's early observational learning studies provided compelling evidence that learning did not require a response contingency and that people could learn by watching someone else be reinforced or punished for their behavior. In the classic "Bobo doll" study by Bandura, Dorothea Ross, and Sheila Ross, Bandura and his colleagues showed that children were, in fact, capable of performing acts of aggression, yet they could inhibit their use of these behaviors. In opposition to behaviorist principles, Bandura's research revealed that people could learn complex patterns of behavior without performing any response or receiving rewards or punishments.

In their 1963 book *Social Learning and Personality Development*, it can be seen that Bandura and Walters retained only a few of Sears's initial learning theory principles and placed increasingly more emphasis on cognitive and information-processing factors. Drawing upon Neil Miller and John Dollard's 1941 book *Social Learning and Imitation*, Bandura and Walters asserted that imitation plays an important role in explaining how novel responses are learned. Imitation in Bandura's early work was conceptualized as a special case of instrumental conditioning, which provided the individual with discriminative stimuli. Another core theoretical assumption stated in Bandura

and Walters's 1963 book was that the environment was believed to contain social cues that would reinforce behavior that matched that of another individual. They hypothesized that imitation could occur even when an individual did not actually reproduce the behavior of another individual during the acquisition phase of learning. In other words, learning could take place in the absence of reinforcement.

Bandura not only challenged psychoanalytic and traditional learning theories, in the late 1960s he began to question the viability of Piaget's stages to explain learning, development, and behavioral changes. Throughout his career, Bandura has adamantly argued that human functioning is much too complex and multidetermined to be categorized into a few discrete stages. Instead of differentiating individuals based on identified stages, Bandura suggested it is more useful to consider the degree to which the individual believes he or she can successfully perform behaviors that are necessary for achieving desired outcomes and the person's expected costs and benefits of performing the activity. The emphasis Bandura placed on self-efficacy beliefs and outcome expectations in his social learning theory were congruent with an increasing interest in cognitive processes among American psychologists. According to Bernard Baars, the "cognitive revolution" in psychology can be traced back to the 1950s in the United States. Cognitive theories and principles were initially presented in Bandura's first presentation of his grand theory in his 1977 book *Social Learning Theory*. He updated his theory in 1986 with the publication of *Social Foundations of Thought and Action: A Social Cognitive Theory*. The change in title from social learning theory to social cognitive theory is significant and represents the increase Bandura placed on the role of cognitive processes in understanding the causes of human behavior, personality development, and thought.

Another contemporary issue that influenced the evolution of Bandura's social learning theory during the 1950s and 1960s was the enormous impact of television, Miller and Dollard's 1941 seminal book *Social Learning and Imitation*, and other forms of mass media in American society. A review of Bandura's work reveals that he has always had an interest in the impact of technological change on the individual and society. For example, some of his earliest research focused on the extent to which violent cartoons affected children's propensity for violent behavior. It is noted that whereas Bandura responded to the impact of

technology, traditional learning theorists, Sigmund Freud, and Jean Piaget developed their theories prior to the television revolution. Social learning theorists believe that technological changes do influence human functioning and development. Moreover, they attribute changes in personality development and human behavior to social factors and cognitive operations such as a person's choice of activities, access to models, social reinforcement and punishment contingencies, personal standards for moral behaviors, societal values, self-regulation abilities, personal and collective goals, and value associated with a specific task.

Bandura's early research focused on aggression, but his research program on modeling quickly evolved to the study of prosocial behaviors. Social learning theorists provided the empirical evidence that children could learn how to share and how to be empathic and altruistic from observing models in their environment.

Two other broad applications of social learning research are in the areas of personality and moral behavior. In terms of personality development, social learning theorists believe that learning that occurs through interactions with other people serves as the primary factor in the development of people's personality. Social learning theorists challenged other psychological perspectives that emphasized moral reasoning and ignored moral behavior.

Core Theoretical Assumptions

Reciprocal Determinism

Social learning theory specifies reciprocal determinism as an essential component of the theory. Reciprocal determinism is the construct referring to a model that represents the idea that human learning and behavior can best be explained by examining the interaction between a person's cognitive processes, behavior, and the environment.

Human Agency

Bandura points out that people "are agents of experiences rather than simply under goers of experience." From the social learning perspective, people are viewed as active players in their everyday functioning and self-development. Bandura distinguishes between three types of human agency: personal, proxy, and collective. Personal agency includes being

able to purposely use information and resources; self-regulate; act on one's beliefs, goals, expectations, values, and expectations; explore and influence environments; and set personal standards for behavior. Having personal agency means that people can exercise freedom, control, self-influence, intentionality, forethought, self-reactiveness, and self-reflectiveness. The exercise of these agentic functions is what enables people to be partial contributors to what they become and do. It is the self-reflective capabilities that shape people's life courses as this type of meta-cognitive thinking allows people to address their motivational conflicts and choose certain actions, personal standards, and environments over others. The ability to reflect upon and make judgments about the merits of one's thoughts and behaviors is what differentiates people from other living beings. Social learning theorists do not attribute human agency to personality or intellectual traits.

In many situations, people rely on the power, influence, competence, and favors of others to meet the demands of everyday life. Proxy agency involves enlisting others to invest their time, effort, and resources so that they mediate and control environments, events, and certain outcomes. Being able to effectively use proxy agency requires the individual to have a good understanding of who in the environment actually has the power and ability to influence the environment. Who people turn to for proxy agency is influenced by the person's self-efficacy for the specific task.

In addition to personal and proxy agency, people rely on collective agency, or the social interdependence of a group, to meet life's challenges. People realize that everyday life is much too complex and demanding for a person to have the control, effort, and resources to manage all situations. They come to understand that some activities and situations are best handled by a coordinated group effort. Social learning theorists believe that people must rely on personal, proxy, and collective agencies in order to adapt and survive.

Unique Human Capabilities

Social learning theorists believe that people are self-regulating, proactive, and self-organizing beings with unique capabilities. The six basic human capabilities are

1. advanced neurophysiological systems,
2. symbolization,

3. forethought,

4. vicarious observation,

5. self-regulation, and

6. self-reflection.

Bandura considered self-reflection to be a characteristic unique to humans. He argued that reflection involves more than making meaning out of one's past and present experiences; it also entails having the capabilities to (a) predict future occurrences, (b) judge the adequacy of one's thoughts based on the consequences related to them, and (c) change one's thought patterns. Self-reflection is considered to be an essential factor in explaining how people self-regulate their behaviors and thought.

Modeling

According to social learning theorists, modeling serves the following three important functions: observational learning, response facilitation, and inhibition and disinhibition. Observational learning is simply the idea that people learn not only from their own personal experiences, but also from watching others. In other words, exposure to a model can lead to learning new behaviors, and a response does not have to be made in order for a person to learn. Thus, observational learning is considered to be "no-trial learning," which runs counter to the prevailing theoretical assumptions in classical and operant learning theories. In social learning theory, the reward and punishment of behavior can occur by watching another person, meaning that a person can experience vicarious consequences to behavior. Observational learning is considered to reflect true learning because people acquire new behaviors through the subprocesses of observational learning (attention, retention, production, and motivation).

The second function of modeling is response facilitation, which is the idea that modeled behaviors serve as social prompts for observers with regard to appropriate behaviors in a given situation. Response facilitation provides the motivational inducements to perform already learned behaviors. Because response facilitation does not lead to the acquisition of new behaviors, it is not considered to reflect true learning.

The third function of modeling is the inhibition or disinhibition of already learned behaviors. When a person watches another person, his or her inhibition to perform certain behaviors can be strengthened or weakened. To inhibit a behavior means to restrain a response or make it occur less frequently. In contrast, disinhibition refers to freeing the behaviors from restraint and allowing them to occur. Disinhibition occurs when an observer sees a model perform a prohibited or threatening behavior without any associated negative consequences.

Learning and Performance

A core social learning principle is that learning and performance can be distinguished from each other. Whether through direct or vicarious experiences, rewards and punishments are expected to affect performance but not learning or the acquisition of new behaviors. It is further assumed that people do not perform everything they learn. The likelihood of people performing what they have learned is influenced by the incentives available in the environment; their evaluative reactions to their own behaviors and the behaviors of others; and the extent to which people consider the pursued activity to be self-satisfying, valuable, and potentially contributable to their self-worth.

Self-Efficacy and Collective Efficacy

In the 1970s, Bandura began to challenge psychological theories, including his own, that ignored the role of self-beliefs in understanding human functioning and development. At present, a substantial amount of empirical evidence reveals that self-efficacy beliefs influence all aspects of people's psychology, learning, and development. The study of self-efficacy has indeed become a central theme among researchers and among social learning theorists in that it serves as the foundation for understanding well-being, personal achievement, and human motivation. Since the publication of Bandura's 1997 journal article "Self-Efficacy: Toward a Unifying Theory of Behavior Change," hundreds of researchers have focused on self-efficacy in such diverse fields as athletics, medicine, education, kinesiology, organizational management, mass media and communication, and many more.

Self-efficacy beliefs play an important role in shaping people's lives by directly affecting people's choices and effort in a given situation. Human functioning is mediated by self-efficacy beliefs, which directly influence people's selection of models and environments. Research shows that people will select certain environments and opportunities and ignore

others. In addition, self-efficacy beliefs affect people's level of effort, which is related to people's evaluation of task difficulty. The strength of self-efficacy beliefs for a given task predicts the likelihood of who will regain their self-assurance following such setbacks. Thus, having a high level of self-efficacy in the face of confusion and failure contributes to people's resiliency and gives them the motivation needed to maintain the effort and to persist in difficult situations. According to Bandura, the relationship between self-efficacy and persistence is mediated through social support selection processes.

Developmental Status of the Learner

Although social learning theorists do not endorse discrete development stages, they do claim that learning is influenced by developmental factors. In addition to lacking certain physical characteristics necessary to perform certain behaviors, children also have less capacity to attend to the behaviors of models for extended periods of time compared to adults. As children develop, they become better at self-regulating their behavior, and the motivational inducements that influence whether or not they will perform what they have learned change as they develop. With time, children are more likely to attend to the behaviors of others and to perform the modeled behaviors that are more aligned with their values and less dependent on immediate reinforcements and punishments of their actions.

Because of developmental limitations, children have difficulty distinguishing between relevant and irrelevant behaviors and verbalizations of a model. Their self-regulation is affected by their limited ability to formulate and maintain well-defined, long-term goals. Moreover, children's performance is affected by the fact that, developmentally, they are not yet capable of ignoring internal and external distractions while they are trying to complete a task. According to Dale Schunk, as children mature, they become more effective at selecting appropriate problem-solving strategies and acting more slowly and deliberately in order to avoid errors. It can be seen over time that children become more proficient in making correct attributions for their successes and failures. They also learn to make better judgments about the level of their own capabilities and are less likely to overestimate their abilities. In educational settings, as children grow, they increasingly base their judgments about their own capabilities by using peer comparisons and rely less on teacher feedback about their capabilities.

Fortuitous Factors in Life

Bandura has criticized theoretical perspectives that do not include fortuitous factors in explanations of human development. He argued that although fortuitous events are unintended, their effects on human development still have some controllability. People shape their lives by using their attributes, belief system, interests, competencies, goals, and personal resources on unplanned events and preparing for how they will respond to chance.

According to social learning theorists, people's accomplishments, status, and success are not caused by luck, chance, good fortune, or good timing. Bandura argued that personal initiative often places people into situations where fortuitous events can affect their lives. Successful people put themselves in certain situations and know a good opportunity when they see it. They know how to capitalize on chance and exploit promising fortuities by cultivating their interests, skills, and enabling self-beliefs so that they are well prepared for the agentic management of fortuity.

Unit of Analysis for Understanding Human Behavior and Thought

Social learning theorists advocate a microanalytic approach to assessing people's goals, self-efficacy beliefs, attributions for successes and failures, and self-regulatory processing. Thus, in order to understand variation in human beings, researchers need to consider social and media factors at the situational level of analysis. The situation-specific approach to explaining human functioning runs counter to trait perspectives and other theories that posit biological determinacy. It is important to point out that a microanalytic approach for studying human behavior and thought does not imply a reductive fragmentary one. Bandura has criticized the field of psychology for the growing interest in deemphasizing psychosocial dynamics in favor of neurodynamics. He warns against fractioning and fragmenting psychology into neuroscience because the study of the whole person and the complex interplay between intrapersonal, biological, interpersonal, and social factors may be lost to conceptual reductionism, nature-nurture dualism debates, and one-sided views about evolution.

Goals and Expectations

Two variables that affect observational learning and performance of learned behaviors are goals and expectations, which social learning theorists argue are best understood and measured in a specific task setting. People who set specific goals are better equipped to engage in self-evaluation based on a standard of performance, which has the potential to increase their motivation, self-efficacy, and learning.

In addition to setting personal goals, people also hold beliefs about the anticipated outcomes of their actions. Based on their previous experiences, and lessons learned from observing models, people form beliefs about the likely consequences of their actions. For example, people who believe they will be successful in a given situation are more likely to stay on task and persevere through difficult times. The beliefs people have about the anticipated outcomes of their actions are reciprocally related to their goals, self-efficacy beliefs, and choice of models.

Conclusion

In conclusion, social learning theorists believe that learning is influenced by how the learner understands the learning event, including the reasons underlying what he or she is doing and why he or she is doing it. At the individual level, early in childhood individuals learn important things, such as rules governing language, from observing others. In adulthood, individuals learn many novel and complex concepts and motor skills through observational learning. Personal factors—such as goals, self-standards, interests, and one's level of inquisitiveness and adventurousness—are known to influence learning and development. Social learning theory is applied to the understanding of individual behaviors as well as issues at a societal level. Research from social learning investigations has furthered social policy and the management of school and workplace social environments. For example, numerous school antiviolence and antibullying prevention programs are designed based on social learning theory principles, especially social modeling and self-regulatory principles. In summary, the triadic model of reciprocal determinism is the central construct in social learning theory and continues to be the causal model used to explain how people's behavior is influenced by the interactions among personal characteristics, behavioral patterns, and environmental factors.

Gypsy M. Denzine

See also Observational Learning; Reciprocal Determinism; Self-Efficacy

Further Readings

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*, 191–215.

Bandura, A. (1977). *Social learning theory.* Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (1986). *Social foundations of thought and action.* Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (1989). Social cognitive theory. In R. Vasta (Ed.), *Annals of child development: Vol. 6. Six theories of child development* (pp. 1–60). Greenwich, CT: JAI.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: Freeman.

Bandura, A., & Rosenthal, T. L. (1966). Vicarious classical conditioning as a functioning of arousal level. *Journal of Personality and Social Psychology*, *3*, 54–62.

Bandura, A., Ross, D., & Ross, S. A. (1961). Transmission of aggression through imitation of aggressive models. *Journal of Abnormal and Social Psychology*, *63*, 575–582.

Holden, G. W., Moncher, M. S., Schinke, S. P., & Barker, K. M. (1990). Self-efficacy of children and adolescents: A meta-analysis. *Psychological Reports*, *66*, 1044–1046.

Stajkovic, A. D., & Luthans, F. (1998). Self-efficacy and work-related performance: A meta-analysis. *Psychological Bulletin*, *124*, 240–261.

Zimmerman, B. J., & Brody, G. H. (1975). Race and modeling influences on the interpersonal play patterns of boys. *Journal of Educational Psychology*, *67*, 474–489.

SPECIAL EDUCATION

At its core, special education is about the individual. In its ideal form, special education provides the structure by which people with exceptionalities may be provided with appropriate adaptations and/or modifications to their environments such that they have the opportunity to reach their individual potential. The field of special education, moving out of its infancy, is still grappling with just how this ideal can be achieved. Researchers are building off of past

successes in an effort to conceptualize new and exciting approaches to the education of individuals with exceptional needs.

The United States is one of the most progressive countries in the world with regard to the enculturation of people with exceptionalities into the fabric of its society. This has not always been the case, and indeed, the United States continues to be challenged by its call to have all of its citizens reach their individual potential. For most of its brief existence, the field of special education has been attempting to find a productive role to address this societal challenge.

Since its conception less than 40 years ago, progress regarding this lofty goal has been debatable. Some may argue that special education programming has not changed fundamentally since the mid-1970s. Others can see the vast and important changes that have taken place, recognizing the abilities of people instead of their disabilities. It is encouraging that special education maintains a place at the forefront of the minds of educational researchers who continue to refine and redefine the important field designed to meet the needs of individuals with exceptionalities.

Historical Backdrop

Civil rights activists of the 1950s and 1960s found in Article 14 of the U.S. Constitution the legal bedrock by which the landmark Civil Rights Act of 1964 was based. In brief, that act stated that it was unlawful to discriminate based on race, gender, national origin, and/or religion. This legislation set the precedent of "equal protection under the law." Disability activists would take their cue from their civil rights brethren. A decade later, after several related pieces of legislation had been enacted, a watershed piece of legislation was enacted that set into motion an educational system that eventually would become known as special education (e.g., the Fair Housing Act of 1968, the Vocational Rehabilitation Act of 1973).

On November 28, 1975, President Gerald Ford signed Public Law 94–142. This legislation was known as the Education for All Handicapped Children Act of 1975. The legislation provided for free and appropriate public education (FAPE) for individuals with exceptionalities to be delivered in the least restrictive environment (LRE). This act, which was amended in 1986 and renamed to the Individuals with Disabilities Education Act (IDEA) in 1990, has since been reauthorized in 1997 and again in 2004.

The Special Education System

Special education is a system designed primarily for individuals between the ages of birth through 21 who have been identified as having certain types of exceptionalities. Under current law, there are 13 categories of exceptionality protected and serviced within this system: specific learning disabilities, speech or language impairments, mental retardation, serious emotional disturbance, hearing impairments, orthopedic impairments, visual impairments, autism, traumatic brain injury, deaf-blind, multiple disabilities, developmental delay, and other health impairments.

Although not identified as a category of exceptionality under federal law, "gifted and talented" is often thought of as an area of exceptionality within the field of special education. Educational programming decisions for individuals with exceptional gifts and talents are usually made at the discretion of individual states and/or localities.

Characteristics of Special Education

Several characteristics of the special education system are designed to maximize individual potential and enhance learning opportunities for individuals with exceptionalities. Those characteristics are individualization, individual instruction, and explicit instruction.

Individualization

Individualization is a student-centered approach to service delivery and instructional decision making that is the hallmark of special education. A team of stakeholders collaborates to develop and implement an appropriate and reasonable program based on the needs and age of each individual.

Individual need is determined through the special education referral, planning, and placement processes. Generally speaking, special education at all levels (i.e., birth to 3 years of age, preschool, and K–12) involves the following five components:

1. Initiation of a referral—could be made by a medical doctor (e.g., at birth), preschool teacher, or K–12 general or special education teacher, among others, when a concern is identified

2. Assessment of individual eligibility (and educational need at the preschool and K–12 levels)—the

type of assessments for eligibility that are used will depend to a great extent on the age and nature of the presenting exceptionality

3. Development of the individualized program—these programs are called different things at different levels, such as Individual Family Service Plan, Birth to 3; Individualized Educational Program (PK exit or graduation); or Individualized Transition Plan (age 15 until appropriate)

4. Determination of the appropriate environment in which services are to be received—hospital, day care, public school, residential facility, and so on (at the K–12 level, this is known as determining the *least restrictive environment*, or LRE)

5. Evaluation of program effectiveness—at all levels, a mechanism by which the individualized program is reviewed and reevaluated as necessary, but is typically done on a yearly basis

Intensive Instruction

Intensive instruction at all levels involves frequent instructional experiences of significant duration related to the individual's needs. It may involve actively engaging individuals in their natural or learning environment by requiring high rates of appropriate response to the material presented, carefully matching instruction to student ability and skill level, providing instructional cues and prompts to support learning and then fading them when appropriate, and/or providing detailed feedback that is directly focused on the task the student is expected to complete.

Explicit Instruction

Individuals with exceptionalities often require more structure and teacher-directed approaches to learning than their nonexceptional peers (the exception often being individuals with exceptional gifts and talents). Three different types of instructional approaches are frequently used when teaching students with exceptionalities. They include basic skills instruction, adaptive instruction, and/or functional life skill instruction.

Basic skills instruction stresses that the individual must learn a specified set of sequenced skills, each a prerequisite to the next. *Adaptive instruction* is designed to fit the demands of a task or setting aligned to a person's needs and abilities. *Functional life skill instruction* teaches students only those skills that will help them succeed in practical matters related to the natural setting, whether it be the classroom, family, or community.

Current Issues in K–12 Special Education

Providing services to individuals with exceptionalities in the LRE has withstood the changes in the special education landscape. There have been many attempts over the years to reframe and rename LRE. Some of the terminology used to describe how to service and educate students with disabilities has been referred to as *mainstreaming*, *Regular Education Initiative*, *inclusion*, and *access to the general education classroom.* Currently, the No Child Left Behind Act of 2001 refers to students and their "access to the general education curriculum." Despite changes in vocabulary, the message is simple and clear. Special education is in a time of tremendous change resulting from legal mandates related to closing the achievement gap.

Numerous research reports have shown that too many students are being placed within the special education system. Additionally, research has shown that there is an overrepresentation of minority students in special education settings. Indeed, productive academic achievement outcomes resulting from special education services and supports arguably have been questionable. As a result, recent legislation has been enacted that attempts to focus educators' efforts on early intervention and prevention.

Promising Practice

There is an emerging body of research that has demonstrated how evidence-based, schoolwide intervention approaches can affect positive academic and behavioral outcomes for K–12 students with exceptionalities. As a result, schoolwide intervention approaches are under rapid development, evaluation, and implementation across the country. The traditional role of the special educator may be changing. Future special educators may act as trainers, consultants, coaches, collaborators, and implementers of specially designed instruction and supports in a schoolwide model, acting as "interventionists." This new type of special educator may have the flexibility and comprehensiveness in training to work across many settings.

Michael P. Alfano

See also Autism Spectrum Disorders; Behavior Disorders; Behavior Modification; Bilingual Education; Gifted and Talented Students; Mental Retardation

Further Readings

Finn, C. E., Rotherham, A. J., & Hokanson, C. R. (2001). *Rethinking special education for a new century.* Washington, DC: Fordham Foundation.

Speech Disorders

Speech refers to the production of speech sounds in a way that conveys meaning in a socially shared system of language. *Disordered speech*, therefore, is the impairment of either the motor production of sounds or the use of the sounds themselves in meaningful contrasts. Respectively, these disorders are typically called *articulation disorders* or *disorders of the phonological system.*

This entry includes information regarding both articulation and phonological disorders in a variety of diagnoses and conditions. Disorders of articulation include not only the developmental delays or disorders commonly heard in typically developing children's speech, but also the disorders that are neurologically or structurally based as well. It is not uncommon, for example, for children with cerebral palsy or with cleft lips or palates to exhibit articulation disorders. Additionally, this entry includes a brief description of the common practices for assessment and intervention of disorders of speech and the phonological system.

Some Key Words

Principle to the discussion of disorders of speech and the phonological system is the concept of intelligibility. *Intelligibility* refers to how easy or difficult it is to understand a person's speech output. In addition to how accurately the speech sounds are produced by the speaker, the communicative context (topic, environment, and power differentials) and listener familiarity with the topic and the speaker also affect intelligibility. Intelligibility is likely to be adversely affected by contexts where the topic is dense, the environment distracting, and the speaker deferential. A listener familiar with the topic or the speaker's particular style can ameliorate the negative effects of the context to some extent. Improvement in intelligibility is thought to be the goal for all interventions for disorders of speech and the phonological system, but it is not always achieved solely by attention to production of specific speech sounds. A speaker can be intelligible yet still exhibit speech errors, if he or she is taught to use strategies to assist his or her listeners.

Speech is made up of speech sounds or phonemes. For the purpose of this entry, these terms can be used interchangeably, despite the nuances of their distinct definitions within the communication sciences community. Therefore, the term *phonemes* will be used exclusively from this point forward. The linguistic function of phonemes is to signal a change in meaning. Hence, when the phoneme /t/ is produced in the word *cat* it distinguishes that word from the word *cap.* The human vocal tract (the pharyngeal, nasal, and oral cavities) is shaped to either impede or direct airflow across the articulators (the tongue, the lips, the teeth, and the hard and soft palates) in order to create a measurable acoustic signal that differs according to the shape of the tract, the presence or absence of laryngeal voicing, and the degree of constriction through which the airflow passes.

Phoneme production varies from individual to individual. For example, the phoneme /s/ is produced in some speakers with the tongue tip pointing downward behind the bottom teeth, whereas others produce it with the tongue tip pointing up behind the top teeth. Regardless of the direction of the tongue tip, the center of the tongue is grooved and the airflow is directed into a very tight constriction in the front of the mouth. The acoustic difference between the resulting productions is not recognized in English as one that distinguishes meaning, however. This type of variation is referred to as allophonic. Each spoken phoneme, then, can truly be conceptualized as a class of acoustic signals rather than a single acceptable production. Hence, the speaker's intelligibility will be affected only if the physical production of the phoneme varies so significantly from the typical that a phoneme is produced that is outside of the class of the target phonemes. An example will help to illustrate this concept.

Many people have listened to a speaker with a dialect different from their own at some point in their lives. In a dialect where the *-ing* ending is reduced to *-in* ("sittin" instead of "sitting"), the ear that is trained to listen to English will not have difficulty

recognizing the representation of the *-ing*. Of course, there are different densities of dialect, and those that are most dense may well impede intelligibility due to crossing of phonemic boundaries. In order for speech to be communicative, it must conform to the standards of the speech community. For the most part, however, dialectal differences are demonstrations of allophonic variations within the phoneme class.

As stated previously, speech disorders can really be separated into two subclasses of disorders. Disorders of articulation are those in which the child is unable to physically produce the speech sound in a way that indicates a change in meaning. There are a variety of error types that may be seen in articulation disorders ranging from frank substitution of a different phoneme to distortions of the correct form that serve to distract from the attempted message. Placing the tongue too far forward in the mouth while producing the phoneme /s/ will result in either a /ʾ/ (the voiceless TH sound) or a "dentalized" /s/. Given that /s/ and its voiced cognate /z/ are quite common in English speech production, these errors are likely to negatively influence intelligibility significantly.

Disorders of the phonological system are those that manifest themselves in a systemic variation in the contrastive rules that make speech communicative. Recall that phonemes vary across the domains of vocal tract shape, laryngeal voicing, and the degree of airflow constriction. If, for example, a speaker does not vary his phonemic productions in the domain of vocal tract shape, the distinction of phonemes produced with an anteriorly placed tongue from the phonemes produced with a posteriorly placed tongue will be lost. This is usually not an indication of a physical inability to place the tongue in the proper position. Instead, what is disordered in this condition is the ability to recognize and represent the linguistic distinction between the resulting phonemic productions.

Articulation Disorders

Just as children are not expected to sit up, stand, walk, or run before the muscle groups supporting such abilities are strong enough and finely enervated enough to do so, a child cannot be expected to produce phonemes accurately before the muscles of the vocal tract have become strong enough and finely enervated enough. Speech production milestones are as predictable as the gross motor skills listed above, with an unvarying progression. Children who are developing

typically can be expected to be intelligible by the time that they are producing sentences (roughly around the age of 4). Recall, however, that intelligibility does not necessarily require perfect production of all speech sounds. The speech errors that typically developing 4-year-olds show may be noticeable, even to an untrained ear, but they do not necessarily impede intelligibility enough to warrant the attention of a speech and language pathologist.

Typically developing children produce approximations to the adult forms of speech as they are developing mastery of the physical forms for speech production. Thus, the children who use /w/ in place of /r/ or /l/ are simply producing a sound that is easier to represent physically than is the adult target sound. There is a variety of literature surrounding the ages at which mastery of particular speech sounds should be expected, but by and large, the conventional wisdom within the communication sciences community is that there is a very wide range of typical development. Most important in this arena, perhaps, is the sequence of phoneme acquisition as opposed to the time line of phoneme acquisition. Simpler sounds, such as /b/, /m/, /p/, /w/, /d/, /n/, and /h/, are expected to be acquired earlier than more complex sounds, such as /t/, /k/, and /g/, and these will be acquired before the more difficult /s/, /z/, /l/, and /r/. Therefore, a typically developing 5-year-old child is not considered to have a disorder of articulation if he is not producing /l/ and /r/ (for example), but is producing all the easier sounds accurately.

Concerns arise, however, in the event that an otherwise typically developing child has errors that persist beyond what is considered the typical developmental period. A 4-year-old who has not yet mastered the production of the early sounds is unlikely to be intelligible in the context of sentence production. In this case, a referral to a speech and language pathologist is indicated.

Other cases of articulation disorders may be attributed to neurological or physical causes, such as cerebral palsy, cleft lip/palate, or hearing impairment. Neurological injury such as cerebral palsy, spina bifida, or traumatic brain injury can result in developmental dysarthria or developmental verbal apraxia. Developmental dysarthria resembles the dysarthrias seen in adults following a cerebral vascular accident (stroke) or traumatic brain injury with the exception that the system that is damaged is one that has not yet fully developed. Children learning to speak using

a vocal tract that is hypotonic, hypertonic, or of varying tone will have a much more difficult time meeting the articulatory postures expected to produce recognizable phonemes. The resulting speech will often be imprecise and of inconsistent vocal quality.

Developmental verbal apraxia, similarly, resembles the apraxias seen in adult onset neurological insults. Although this diagnosis can be assigned accurately to many children who clearly have intent to communicate but lack the motor control to form speech sounds accurately, it has been applied somewhat capriciously to other children who are nonverbal but have no known neurological history. A child with an accurate diagnosis of developmental verbal apraxia is likely to be demonstrating not only difficulty in producing speech, but other subtle neurological signs as well. A thorough diagnostic battery for developmental apraxia of speech ought to include neuropsychological and in-depth developmental testing as well.

Children with cleft palates will show different types of articulation errors. Depending on the extent and severity of the cleft, the child may have difficulty closing off the nasal cavity in order to produce sounds that resonate in the oral cavity (the majority of sounds in the English language). Additionally, the child with a cleft palate may have difficulty building up enough pressure in the oral cavity to produce some of the sounds that require a tightly constricted airflow. Therefore, children with cleft palates frequently exhibit sharply reduced intelligibility because of the hypernasality of their speech and the imprecision of their consonant production. Speech and language pathologists working with children who have cleft palates will usually rely on a team approach, consulting with oral and plastic surgeons, otolaryngologists, audiologists, and counselors. Similar difficulties will arise in children with cleft lips only, but they tend to be of lesser severity and will not always require the intensity of intervention seen with children with cleft palates.

Finally, children with hearing impairment have limited ability to self-monitor their speech sound productions because of their auditory limitations. The advent of the cochlear implant and its widespread use with children as young as 6 months of age will likely improve the speech outcomes of many children with hearing impairment over the course of the next decade. Still, not all children with hearing impairment are candidates for cochlear implants, and therefore, educational and health professionals will continue to see children with hearing impairment and articulation disorders. Speech and language pathologists are developing programs to address the specific needs of hearing-impaired children who strive to become oral communicators. Within the domain of hearing impairment, it is important to also consider the fluctuating hearing impairment that co-occurs with chronic otitis media (middle ear infection). The children who have fluctuating hearing losses typically struggle not with the motor aspect of speech, but the establishment and application of the phonological rules necessary to use the speech sounds accurately in communication.

Phonological Disorders

The phonological system develops predictably as well. Children first learn to distinguish between the phoneme classes of vowels and consonants (vowels require very little airflow restriction, consonants require degrees of constriction), then develop the distinctions between consonants that actually impede the airflow and sounds that merely constrict the airflow, and so on. This progression follows a predictable developmental pattern that is not only influenced by the ambient language surrounding the child, but can assist the adults in the child's environment in interpreting his or her earliest attempts at speech. Children who are typically developing are frequently heard to substitute whole classes of sounds with other classes of sounds in what the communication sciences community calls *phonological processes.* For example, children who are developing typically frequently substitute "stop" consonants (e.g., /p/, /b/, /t/, /d/) for "continuant" consonants (/f/, /v/, /s/, /z/). Thus, a child attempting to say "We saw a fat fish!" might produce something more like, "We taw a pat pit!" Some additional examples of typical phonological processes include the deletion of final consonants (the /t/ in *cat*) in words, substitution of front consonants (/t/ and /d/) for back consonants (/k/ and /g/), and substitution of glide consonants (/w/ and /j/) for liquid consonants (/r/ and /l/). There are expected ages at which each of these phonological processes becomes suppressed, and the children who persist in using these processes or who use processes that are atypical are considered to be using a disordered phonological system.

It is important to note that the two types of speech disorders do not always occur in exclusion to one another. Children with cleft palate, for example, can exhibit the motor difficulties in producing the sounds

and also exhibit a persistent process of final consonant deletion.

Additionally, it is not uncommon for either speech disorder or for both speech disorders to be present in conjunction with other developmental disorders. Children who have language-learning disorders (e.g., dyslexia) are frequently also diagnosed with disorders of speech and/or the phonological system, especially in the areas that require automaticity in the application of phonological rule systems (e.g., spelling, phonemic awareness). Children with various syndromes frequently present with disorders of either type of speech disorder as well.

Assessment for Speech Disorders

A child who is suspected of having a speech disorder can be referred to a speech and language pathologist. Speech and language pathologists begin by screening a child's speech to determine if the child's speech is developing typically or if further evaluation is necessary. Speech screening is usually completed in a very brief encounter in which the child is asked to say words in which the most commonly misarticulated sounds are represented. The stimulus materials for a speech screening generally range from imitation to picture-naming tasks. If the child's speech is determined not to be developing typically, a full evaluation is then indicated. The tools predominantly used to screen hearing are norm-referenced and provide specific cutoff scores for determining paths of involvement following the screening.

A full evaluation will include listening to the child produce sounds in single words, in sentences, and in continuous speech. Single words are typically elicited via picture-naming tasks, sentences via story re-tell or extended imitation, and continuous speech via a structured conversation. Both criterion-referenced and norm-referenced materials are available and may be used for this stage of the assessment. The child's productions are transcribed using the International Phonetic Alphabet and the diacritical marks that indicate the types of distortions that may be present in the child's speech. Although single words are easiest to transcribe and analyze, an evaluation is not considered to be complete until a speech sample of continuous speech has been analyzed as well. Recall the discussion concerning intelligibility—it is conceivable that children who can produce speech sounds without error in single word contexts may be far less than

intelligible in continuous speech. Therefore, the additional diagnostic information gained from a continuous speech sample justifies the additional time and effort spent on transcribing in that context.

Once error sounds are identified, the speech and language pathologist analyzes them to determine patterns of production and substitution. Errors that appear to be due, principally, to deviations in motor production are addressed through traditional articulation methods, which focus on teaching proper placement for production of the sounds and intensive practice of correct production in hierarchical contexts, from isolation through continuous speech. Patterns of errors across or within phoneme classes, however, indicate a disorder of the phonological system and require intervention that focuses on teaching the linguistic function of the phoneme classes.

Once errors are classified, the speech and language pathologist will generally assess stimulability, or the child's ability to match the adult's model of correct phoneme production. Stimulability can be assessed in isolation (producing a phoneme in a single syllable context), single words, or sentences. Traditionally, the child is told to watch the speech and language pathologist's mouth and to listen to the correct production, then try to match it. The results of stimulability testing are used to determine appropriate goals for intervention and the methods that are best suited to the child's particular need.

Intervention for Speech Disorders

The motor aspect of speech disorders lends itself to myriad comparisons to other gross and fine motor skills. However, the motor aspect must be considered and addressed in the context of authentic communication if the child is going to achieve intelligible speech. Mastery of the motor production of isolated phonetic positions will not allow the child to produce intelligible, continuous speech unless the child has the opportunity to practice that production in that context. One area of accurate parallel between other motor skills and speech, however, is the need for frequent and intensive practice (in the most advanced context the child can master) in order to develop the proper habitual placement of the articulators that produce the target sound accurately.

Determining which type of disorder the child is manifesting leads to the formulation of goals for intervention. Much of the methodology for intervention of speech disorders follows a behavior modification

theme. As stated above, the child who is manifesting a disorder of motor production only is taught the proper placement and provided with frequent opportunities to practice the target sound in hierarchically arranged contexts (shaping). The traditional articulation method also includes training in auditory discrimination between the target production and error production. This training provides the child the opportunity to learn to distinguish between his or her habitual error production and the correct production. Without this skill, training in correct production may not be as successful because of the child's inability to determine the difference in sounds produced by correct and incorrect placement.

The hierarchical order of contexts begins with production of the sound in isolation or with a vowel to form a syllable. The child and the speech and language pathologist practice at this level until the child can produce the sound accurately and consistently, and then the context is advanced to the level of single-word production. Once again, the child and the speech and language pathologist practice at this level frequently until the child produces the target production accurately and consistently. The contexts continue to advance until the child is producing the target sound in continuous speech accurately and consistently.

If the child's speech disorder is phonological in nature, the focus of intervention becomes teaching the contrastive function of the phonemes. Speech and language pathologists who use this method to treat phonological speech disorders must assume that the child's motor function is without error. If this assumption cannot be met, motor training must precede, or at least coincide with, the phonological intervention. The speech and language pathologist addresses phonological system errors using minimal pairs, words that vary by a single phoneme (e.g., *cat* and *cap, Kay* and *cave, Wayne* and *rain*). Explicit teaching of the need to contrast words via correct phoneme usage precedes the practice with the minimal pairs.

Within the practice of intervention for speech disorders, there may be a need to address alternative or augmentative communication due to the severity of the intractable nature of the speech disorder. At that time, the speech and language pathologist collaborates with specialists in assistive technology in designing a device that will serve the child's communicative needs in his or her ambient environment.

Martha Dunkelberger

See also Assistive Technology; Behavior Modification; Communication Disorders; Criterion-Referenced Testing; Dyslexia; Language Disorders; Learning Disabilities; Norm-Referenced Tests; Shaping; Spelling

Further Readings

Bauman-Waengler, J. (2004). *Articulatory and phonological impairments: A clinical focus* (2nd ed.). Boston: Pearson Education.
Flipsen, P. (2003). Articulation rate and speech-sound normalization failure. *Journal of Speech, Language and Hearing Research*, *46*, 724–737.
Flipsen, P., Hammer, J. B., & Yost, K. M. (2005). Measuring severity of involvement in speech delay: Segmental and whole-word measures. *American Journal of Speech Language Pathology*, *14*, 298–312.
Lewis, B. A., & Freebairn, L. (1998). Speech production skills of nuclear family members of children with phonology disorders. *Language and Speech*, *41*, 45–61.
Lewis, B. A., Freebairn, L., Hansen, A., Stein, C. M., Shriberg, L. D., Iyengar, S. K., et al. (2006). Dimensions of early speech sound disorders: A factor analytic study. *Journal of Communication Disorders*, *39*, 139–157.
Lewis, B. A., Freebairn, L., & Taylor, H. G. (2000). Follow-up of children with early expressive phonology disorders. *Journal of Learning Disabilities*, *33*, 433–444.
Owens, R. E. (2007). *Language development* (7th ed.). Boston: Pearson Education.
Shriberg, L. D., & Kent, R. D. (2003). *Clinical phonetics* (3rd ed.). Boston: Pearson Education.
Shriberg, L. D., Kent, R. D., Karlsson, H. B., McSweeny, J. L., Nadler, C. J., & Brown, R. L. (2003). A diagnostic marker for speech delay associated with otitis media with effusion: Backing of obstruents. *Clinical Linguistics and Phonetics*, *17*, 529–547.
Shriberg, L. D., Tomblin, J. B., & McSweeny, J. L. (1999). Prevalence of speech delay in 6-year-old children and comorbidity with language impairment. *Journal of Speech, Language and Hearing Research*, *42*, 1461–1481.
Sutherland, D., & Gillon, G. T. (2005). Assessment of phonological representations in children with speech impairment. *Language, Speech, and Hearing Services in Schools*, *36*, 294–307.

SPELLING

Success in modern literate societies requires the ability to read and write proficiently. Although the past several decades have seen a strong interest in the

processes underlying proficient reading ability, much less research attention has been given to understanding writing. This is due in part to the fact that writing is a historically more newly acquired skill than reading and, until recently, was a skill that was mastered by few people. However, with the burgeoning of technological advances, success in many societies now requires strong literacy skills that include proficient writing skills. An important component of the writing process is spelling. The ability to produce written words easily and proficiently allows the writer to have sufficient cognitive resources available for other important aspects of writing, such as composing the message. As such, researchers have increasingly begun to focus on the structure of the spelling system and the development of spelling abilities.

Most theories of the structure of the spelling system assume that there are two independent processes responsible for the spelling of familiar and unfamiliar words: a lexical process, which is mainly responsible for spelling familiar words, and a sublexical process for spelling novel or unfamiliar words (see Figure 1). When spelling a familiar word via the lexical process, the spelling is retrieved from a long-term memory store called the orthographic lexicon. The forms in this store comprise graphemes that are abstract representations of letters that can be converted into either written or oral form in later stages. The lexical process is semantically mediated such that in order to spell a familiar word, a stored spelling is retrieved from the orthographic lexicon via semantic input, which includes information about the meaning of the word to be spelled. The abstract orthographic representation that is retrieved is kept active via a working memory process called the graphemic buffer. The buffer is responsible for keeping the orthographic information active so that serial processes can convert the abstract form into an appropriate oral (letter names) or written form (letter shapes) for output. When an unfamiliar word must be spelled, the sublexical process serially converts the oral input, which comprises phonemes (sounds), into a string of abstract letter representations (graphemes). These are then kept active by the graphemic buffer for serial output for either written or oral spelling. The sound-to-letter conversion process is guided by the frequency of phoneme-to-grapheme mappings, which indicate which sound-letter mappings are most common, as well as by contextual constraints that indicate which letter combinations are permissible and most common in a language. Recent evidence

suggests that lexical and sublexical processes interact during spelling such that sublexical processes influence the output of the lexical system.

Early theories of spelling assumed that the spelling representations stored in the orthographic lexicon are linear strings containing only information about letter identities and their order. However, studies of patients with impairments to the graphemic buffer have provided evidence that orthographic representations are more complex than originally proposed. Damage at the level of the buffering process should result in errors that reflect the limitations of a working memory process: errors reflecting differences in length and letter errors such as misordering letters, deleting letters, and so on. However, some patients have produced distinctive patterns of errors that indicate that orthographic representations include not just letter identity and order, but also the consonant-vowel status of the individual letters and information about letter quantity (i.e., whether a letter is doubled). More controversially, some data suggest that information about the written syllable structure of a word is also represented. Overall, these data indicate that the long-term memory representations of words' spellings are complex, including information that goes beyond a simple listing of letters and their order.

When examining the development of spelling skill, several abilities form the foundation for becoming a proficient speller. As with reading, children must first learn that writing represents spoken language. After learning this concept, children must learn to segment speech into the linguistic units of the language, whether those units are syllables of syllabary languages or phonemes in alphabetic languages such as English. Another important task is learning the symbols that are used to represent these linguistic units. Although much research has focused on alphabetic languages that use letters to represent phonemes, many languages use different systems for representing speech in writing, such as Chinese, which uses logographic symbols to represent morphemes, which are the smallest meaningful units in a language, such as root words and affixes. The task of learning the written symbols that represent speech can tax memory resources. Adding to the difficulty is the fact that relationships between written symbols' shapes and sounds and their shapes and names are typically arbitrary. Furthermore, once children have learned to segment speech and recognize the written units of representation, they must also learn to produce
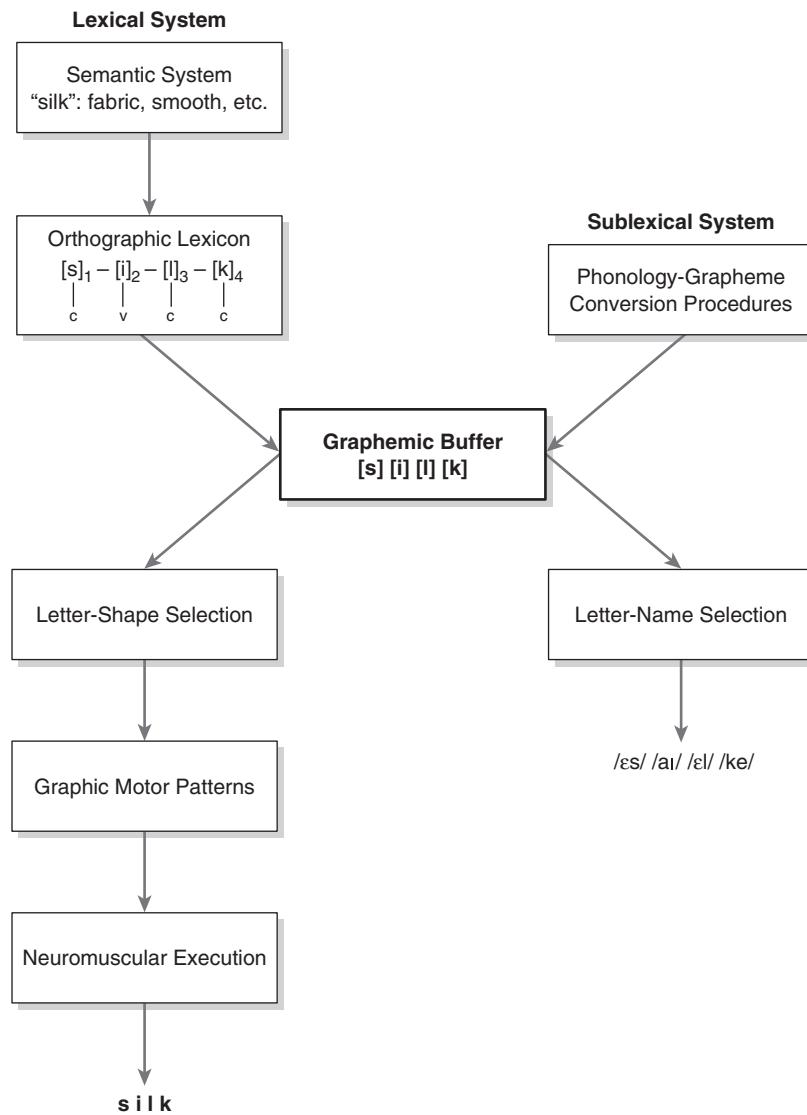
Figure 1   A Dual-Process Model of the Spelling System

the written symbols. Research in several alphabetic languages suggests that phonemic awareness skills are related to spelling ability in children.

Recent work suggests that children learning to spell English do not simply memorize sound-letter correspondences. Rather, they use their knowledge of letter names, which often are related to their sounds, to aid in spelling. Additionally, children are more accurate spellers when they consider the context in which a unit occurs. However, for a child to extract associations requires exposure to enough words to realize that certain linguistic units can be spelled multiple ways and that some spellings are more likely in

certain conditions. Thus, with enough exposure to print, children become aware of the statistical properties of their language and use these properties when spelling familiar and novel words.

Jocelyn R. Folk

See also Dyslexia; Reading Comprehension Strategies

Further Readings

Miceli, G., & Capasso, R. (2006). Spelling and dysgraphia. *Cognitive Neuropsychology*, *23*, 110–134.

Treiman, R., & Kessler, B. (2005). Writing systems and spelling development. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 120–134). Oxford, UK: Blackwell.

STANDARD DEVIATION AND VARIANCE

The standard deviation (abbreviated as *s* or *SD*) is the average amount of variability that a set of scores contains and is the average distance from the mean. The larger the standard deviation, the more variability is in the set of data and the farther the average point is from the mean of the set of scores.

The formula for computing the standard deviation is as follows:

$$s = \sqrt{\frac{\Sigma(X - \overline{X})^2}{n - 1}},$$

where

s is the standard deviation

Σ is sigma, which tells you to find the sum of what follows

X is each individual score

$\overline{X}$ is the mean of all the scores

n is the sample size.

The variance is simply the square of the standard deviation.

Given the following data set,

56

42

44

67

78

76

56

45

65

61

the steps for computing the standard deviation are as follows. See Table 1 for the corresponding value for each step.

1. List each score.

2. Compute the mean of the group (in this example, it is 59).

3. Subtract the mean from each score.

4. Square each individual difference. The result is the column marked $(X - \overline{X})^2$.

5. Sum all of the squared deviations about the mean (in this example, it is 1,482).

6. Divide the sum by $n - 1$.

7. Compute the square root. That is, the standard deviation for this set of 10 scores is 12.83. The variance would be this value squared or 164.61.

Given these results, each score in this set of 10 differs from the mean by an average of 12.83 points.

The value of 1 is subtracted from the denominator of the standard deviation formula because *s* is an estimate of the population standard deviation and is unbiased. By subtracting 1 from the denominator, the standard deviation is forced to be larger than it would be otherwise. This is done to ensure that if any error

Table 1 Corresponding Values

X	$X - \overline{X}$	$(X - \overline{X})^2$
56	−3	9
42	−17	289
44	−15	225
67	8	64
78	19	361
76	17	289
56	−3	9
45	−14	196
65	6	36
61	2	4
	Total	1,482

Note: M = 59, *SD* = 12.83.

is made, it will be an overestimate of the population value.

Neil J. Salkind

See also Mean

Further Readings

Salkind, N. J. (2003). *Statistics for people who (think they) hate statistics.* Thousand Oaks, CA: Sage.

STANDARDIZED TESTS

A standardized test is one that is developed to maximize the comparability of scores by providing all examinees with the same (or parallel) content that is administered and scored in a consistent manner. The term *standardized test* is frequently used incorrectly as a synonym for a multiple-choice test, norm-referenced test, or commercially developed test. Standardized tests can be norm-referenced, criterion-referenced, or standards-based. They can contain one or more of a variety of different item types.

Standardization is not a binary concept. Tests can standardize fewer or more of the conditions of development, administration, and scoring and provide more or less specificity for each condition.

History

In approximately 2200 B.C.E., a formal system of civil service examinations was begun in China. There is no record of the content or methods used. In 1115 B.C.E., content domains were standardized to include music, archery, horsemanship, writing, and arithmetic. Over the years, the content domains changed somewhat, and in 606 C.E., the content and administration methods were further standardized into a system called *Keju.* The Keju system had three levels of competition: "Budding Geniuses," "Promoted Scholars," and "Ready for Office." An example of the degree of standardization in this system is that for the "Promoted Scholars" competition, candidate essays were rewritten by a scribe and marked with a code so that examiners would not be able to recognize the author nor have their judgments affected by penmanship (although penmanship was judged explicitly for the "Budding Geniuses" competition).

More recent roots of standardized testing in the United States stem from Horace Mann's work as superintendent of schools for Boston, Massachusetts, around 1845, when he pushed to replace the tradition of oral examinations with essay testing. Mann compared the earlier, less standardized approaches to running a cross-country race where each runner was timed on the mile he ran and then the next runner started. Each runner would be subject to different conditions, some running on level ground, others running up hill, and others slogging through the mud.

Standardized essay testing became more and more popular at the beginning of the 20th century; however, there was a growing concern over the variations in scoring by different graders. E. L. Thorndike and his students had invented a number of objectively scored item types, including a precursor of the multiple-choice item for use in psychophysical research.

In 1915, Frederick Kelly pursued the development of objectively scored tests that reduced the time and effort for the administration and scoring of reading tests. His criteria for such items were that they should be subject to only one interpretation; call for only one thing; and be wholly right or wholly wrong, and not partly right and partly wrong. Following is the first known published selected response item, which appeared as a practice item on the Kansas Silent Reading Test. Other item types, particularly short answers, were used on this test.

> Below are given the names of four animals. Draw a line around the name of each animal that is useful on the farm:
>
> Cow
>
> Tiger
>
> Rat
>
> Wolf

Frederick's invention influenced Arthur Otis, who used it for about half the questions in his "Scale for the Group Measurement of Intelligence." In 1917, Otis was part of the group of psychologists called together to help the army deal with the problem of quickly and cost-effectively classifying (for training purposes) 1.7 million draftees. Their solution was the first all-selected-response test, the Army Alpha.

The effect of the success of the Army Alpha and its ability to standardize scoring as well as the conditions of administration was to greatly increase the role of standardized testing in the United States. In 1926, the College Board added multiple-choice items to the SAT. The GI bill in 1944, and the anticipation of American colleges needing to test and enroll large numbers of returning servicemen, led to the increased use of multiple-choice standardized tests for college admissions.

In 1937, IBM produced and sold a scanner that sensed graphite marks on paper by determining changes in electrical conductivity. This increased the efficiency of scoring multiple-choice items. This efficiency was further increased by E. F. Lindquist's invention of the optical test scanner in 1955.

Although non-multiple-choice standardized tests (or tests that have both multiple-choice and constructed response items) have remained common (including the Advanced Placement Examination, New York State Regent Examinations, and many others), the dominant form of standardized testing has been the multiple-choice examination.

Item Types Commonly Used in Standardized Tests

There are two families of item types: selected response and constructed response.

Selected Response

Selected response items require examinees to identify one or more correct responses from the choices provided by the test publisher. The most commonly used selected response item type in standardized tests is multiple choice. Typically, examinees use a pencil to fill in an outlined space (often referred to as a bubble, although it might be rectangular, ovoid, or circular) indicating the best response from the several that are provided. Research on the efficiency of the testing process (amount of information gathered per unit time) strongly suggests that three choices would be optimal, although most multiple-choice items provide four or five choices. On computer-administered versions of multiple-choice items, the examinee might click on a circle or on the text or picture of the best response.

Other common selected response item types include true-false, yes-no, ordering, and matching.

Constructed Response

Constructed response items require examinees to create a response to a question. That response might be short (a word, phrase, number, or mathematical expression); long (an essay or a worked-out solution to a problem); or in between. Some constructed response items could arguably be called extended selected response items. That is, the examinee chooses the correct response from an extremely large set of possibilities. One such example is the gridded response item.

The gridded response item type allows examinees to bubble in a numeric answer. Rather than choosing among a small number of provided responses, the examinee would choose from the integers and a decimal point to make up a numeric response, such as 123.45. Some gridded response items also allow the bubbling of arithmetic operators to indicate fractions or formulas, such as $(3/4) * X^2$. Gridded response items have several advantages compared to multiple-choice items. They cannot readily be solved by guessing or back-solving (plugging each possible answer of a math problem into the equation to see if the answer works). Also, their solution paths more closely parallel the problems students will encounter in real life and learn to solve in class. In addition, they are less influenced by test wiseness.

The use of computers has led to a proliferation of constructed response item types, such as drag-and-drop items, which could be used for questions such as placing chemical elements into their proper position in the periodic table.

Scoring

Standardizing the Scoring of Multiple-Choice Items

The proliferation of multiple-choice testing was driven by the relative standardization of the scoring process compared to that of constructed response items. Originally, multiple-choice items were scored against a typed answer key by human graders, but human graders varied in their ability to apply scoring keys accurately. The scoring process for multiple-choice items was further standardized by the invention and use of scoring stencils—pieces of cardboard that had cutouts in the spaces where the correct answers should have been indicated by the examinee. If graders saw a proper mark through the cutout, the examinee was given credit. If graders saw no mark,

they indicated the response's incorrectness by making a red mark. The graders then counted up the marks and indicated the score. To further help ensure accurate, consistent grading (that is, to increase standardization), light tables were used. These devices made light marks made by examinees more visible so test scores were less affected by the visual acuity of the graders.

Hand scoring was slow and could still vary slightly depending on the grader. Electronic and then optical scanners increased the standardization of the scanning process. But scanners, too, were susceptible to variation due to the accuracy of placement in the printing of the bubbles; the moisture content of the paper; the temperature of the scanning room; and the extent to which examinees followed directions carefully, filled out the intended bubble completely, and did not make stray marks. Most recently, companies that scan test answer sheets have created software that intelligently processes the scanning data to try and adjust for any lack of standardization in the aforementioned areas.

Standardizing the Human Grading of Constructed Response Items

The early challenge of scoring constructed response items was that each grader applied her or his own standards, and those standards varied significantly among graders. Thus, students' scores were affected by the luck of the draw in who scored their papers. Standardization of the scoring of constructed response items has depended on the use of scoring guides (also called rubrics), exemplar papers (a set of examples for each score point that helps to further define those score points), and training. It also is best served by the creation and selection of constructed response items amenable to consistent scoring, including field testing the scorability of items.

Scoring guides (rubrics) provide direction to graders. They indicate what an examinee's response must include to achieve a certain number of points (most constructed response items are scored in a nonbinary fashion, often allowing between zero and as many as six points for a response).

Exemplars (sometimes called benchmarks) are often used to provide graders with specific examples for each score point. Exemplars may be annotated to point out what aspects described in the scoring guide were included to justify the score and what aspects were missing that prevented a higher score.

Training provides greater standardization by allowing the trainer an opportunity to provide feedback to graders who grade too harshly or too leniently, or who attend to aspects of the response not intended by the scoring guide. Often, training is followed by a qualifications test where graders must demonstrate their ability to score accurately.

Ongoing quality assurance also helps improve the standardization of the scoring process. There are three common forms: double blinds, read behinds, and consensus papers (other names are used for these three methods—they have no commonly agreed-upon names). Double blinds require a subset of papers to be randomly assigned to a second grader without the second grader knowing the score assigned by the first grader (hence the name *double blind*).

In addition, or alternatively, a sample of each grader's papers might be reviewed by a senior grader. In this case, it is typical for the senior grader to know the score that the first grader provided, although arguably it is better if the senior grader first scores the paper and then looks at the original score.

Consensus papers are usually a small set of papers representing all score points that were scored by a committee of experts (usually as part of the picking of exemplar papers). These papers are interspersed among the work being operationally scored. Consensus papers allow the accuracy of each grader to be compared in a consistent fashion.

With any of these methods, if a grader is found to be regularly scoring papers more harshly or leniently compared to the criterion grader, a more senior grader who has previously demonstrated a high level of accuracy reviews the work of the first grader and provides additional feedback or training, or else removes that grader from the scoring process.

Computer Grading of Constructed Response Items

Short answer questions are fairly amenable to scoring by computer if the answers have been entered by the student in electronic form. Although it is also possible to transcribe handwritten student responses, this can be expensive, and if transcription errors cause an examinee to fail a high-stakes test, there is significant potential risk to the test's sponsor and testing contractor.

Computer grading of short answer questions with numerical answers has often allowed either a list or a range of answers. For example, if the answer to

a problem was 1/6, answers of .17, .167, and .1667 might also be considered acceptable. Because the number of decimal places indicated by students might vary considerably, an alternative might be to consider acceptable any answer that, when represented by a decimal, is between .16666666 and .17.

A problem that requires a word as a response might be scored against a list of key words, including synonyms, homonyms, and common misspellings (unless it is a spelling or word usage test). Other than gridded response items (discussed previously), few major tests currently use computers to grade short answer items.

Research on the use of computers to score essays has gone on since Ellis Page's work starting in 1964. Much recent research has led to a proliferation of scoring systems to score writing samples. In most cases, these systems are first "trained" by inputting samples of papers that have been scored by expert human graders. The purpose of this training is to allow the system to discern the writing features that are associated with each score point. The specific features looked at are typically considered proprietary by the software developers. Often, these systems are said to use artificial intelligence, but given the black box nature of the software and the results of some small-scale studies, it is best to think of them as *predicting* (with reasonable accuracy) what score a human grader would assign. In fact, several computer scoring systems have demonstrated that, after training with a sufficiently large sample of papers, the correlation in a second independent sample between the scores assigned by human graders and the computer is higher than the correlation between two sets of human graders.

Responses that are significantly dissimilar to the papers in the training set are less likely to predict accurately the score a human grader would have assigned. Also, some educators and researchers have worried whether such scoring engines might be easy to fool, perhaps by interspersing long common words or increasing sentence length, even when these actions are uncalled for or nonsensical.

Until very recently, little research had been done regarding computer-based scoring of essay questions that require substantive responses (for example, "Discuss the primary political and social factors that led to the American Civil War"). Such questions are much more difficult to score using computers than either short answer questions or writing quality because the same key words and phrases might be used to argue for or against a point, making it hard to assess the correctness of a response.

Score Interpretation Approaches

A score by itself conveys little meaning. A test can be made of easy or hard items, so knowing the number (or percentage) of items answered correctly provides little insight regarding how well an examinee performed. Developers of standardized testing programs have created and refined a variety of ways to enhance the interpretability and utility of test scores. Four common interpretation approaches are Normative, Ipsative, Criterion-Referenced, and Standards-Based.

Normative Interpretation

Normative interpretations describe examinee performance by comparing an examinee's score with the scores from a meaningful reference group. For example, a high school student might be said to have scored better than 95% of all ninth-grade students in U.S. public schools, or an applicant to graduate school might have scored better than 60% of all graduate school applicants to psychology programs.

Ipsative Interpretation

Rather than compare an examinee to other examinees on the same measure, ipsative interpretations compare an examinee's standing on one measure with that same examinee's standing on other measures. Ipsative interpretation requires that a set of measures be developed so that scores across the set sum to a constant (that is, you cannot have very high or very low scores on every measure). This is done by forcing an examinee to choose among alternatives that score positively on different scales. For example, a question might ask, "Would you rather go to a party or manage a project at work?" An ipsative approach is most frequently used for interest and personality measures. An ipsative interpretation might say, for example, that an examinee's need for dominance is stronger than his or her need for affiliation.

In educational testing, the term *ipsative assessment* is sometimes used to describe comparing a student's achievement over time.

Criterion-Referenced Interpretation

Criterion-referenced interpretations compare examinee performance with a fixed standard of success.

Criterion-referenced tests require the determination of cut scores that define how good of a performance is good enough. Although by definition, only 50% of examinees will score above the 50th percentile (a norm-referenced interpretation), regardless of where a cut score is set on a criterion-referenced test, it is possible for all students to exceed (or fail to exceed) that score.

Standards-Based Interpretation

Standards-based interpretation is a particular type of criterion-referenced interpretation that is supported by content standards that are used to drive both the curriculum and the assessment. It answers the question "What should students know?" and provides performance standards—multiple cut scores that differentiate level of performance and answer the question "How well do students understand the content?"

Overstandardization

As stated earlier, the intent of standardization is to maximize the comparability of scores by making the conditions of the testing enterprise as similar as possible for all examinees. More than 100 years of research shows that standardization works to achieve this goal. But research also shows that inappropriate (or too much) standardization can also interfere with this goal.

Consider an extreme example. We want to administer a multiple-choice high school geometry test. In the name of standardization, we decide that all students must respond with their right hands. Yes, we have increased standardization, but left-handed students are slower and less accurate filling in bubbles with their right hands than they are with their left. And what about students who have no right hand?

Standardization is only useful when it reduces construct-irrelevant variance (in this case, any variability in scores not associated with the examinees' knowledge of and ability to apply geometry). When it instead increases construct irrelevant variance, standardization becomes problematic.

Accommodations

One issue with finding the balance between too much standardization and not enough is that for a subset of the examinee population, a particular standardization decreases irrelevant variance but for another subset, it increases it. Moreover, it can be time consuming and expensive to determine accurately who is in which group. Nonetheless, many testing programs do try to make these distinctions and determine whether certain examinees should be afforded particular accommodations during the test administration.

Some of these accommodations are obvious and noncontroversial, such as allowing a blind person to respond to a test form printed in Braille. The most common accommodation, extra time, is usually provided for students with diagnosed reading disabilities or attention deficits, but studies have shown that many students who are not so diagnosed, and thus not afforded extra time, would have increased their scores were they granted extra time. There are numerous other accommodations that some standardized testing programs allow. Some examples include individual administration (to reduce distractions), breaking a test into multiple short administration blocks, verification of understanding of directions, use of a calculator, translation into a student's native language, use of a bilingual dictionary, sign language, large print, computer entry of essays, and student dictation of answers to a scribe.

Neal Kingston

See also Aptitude Tests; Criterion-Referenced Testing; Grade-Equivalent Scores; Grading; Multiple-Choice Tests; Norm-Referenced Tests; Rubrics

Further Readings

Bejar, I. I., Williamson, D. M., & Mislevy, R. J. (2006). Human scoring. In D. M. Williamson, R. J. Mislevy, & I. I. Bejar (Eds.), *Automated scoring of complex tasks in computer-based testing.* Mahwah, NJ: Lawrence Erlbaum.

Cohen, A. S., & Wollack, J. A. (2006). Test administration, security, scoring, and reporting. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 355–386). Westport, CT: Praeger.

Millman, J., & Greene, J. (1988). The specification and development of tests of achievement and ability. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 335–366). New York: American Council on Education.

Schmeiser, C. B., & Welch, C. J. (2006). Test development. In R. L. Brennan (Ed.), *Educational measurement.* Westport, CT: Praeger.

STANDARD SCORES

Suppose you took a math test and learned that your score was a 47. How would you know if that was a "good"

score? One of the things you might want to know is the range of scores that was possible for the test. Most people would consider a score of 47 on a test with a maximum possible score of 50 better than a 47 on a test with a maximum possible score of 100. A different approach would be to see how your score compared to that of others. For a test on which scores of 0–80 were possible, you might feel better about a score of 47 if you learned that the average score was 42 than if it was 62. *Standard scores* (sometimes referred to as *scaled scores*) provide a means of interpreting a score in terms of its distance to the average score and takes into account the overall variability or spread of the set of scores.

More specifically, standard scores indicate the location of a score relative to the mean of all scores in standard deviation units. *Mean* (*M*) is the mathematical term for the numerical average found by adding all of the scores (often represented as *X*) and dividing by the number of scores. It is one of three measures of *central tendency* used in describing sets of scores. *Standard deviation* (*SD*) is a measure of the *spread* or *variability* of a set of scores based on the distance between each score and the mean of the set of scores: The greater the *SD*, the more spread out the scores are.

Although standard scores take many forms, they are all based on *z scores*. A *z* score is a standard score with a mean of 0 and a standard deviation of 1. *Z* scores are found by subtracting the mean from the score and dividing by the standard deviation:

$$z = \frac{X - M}{SD}.$$

For example, for a set of scores with a mean of 60 and a standard deviation of 8, the *z* score for a score of 60 would be 0, and the *z* score for a score of 76 would be

$$z = \frac{76 - 60}{8} = \frac{16}{8} = 2.$$

For this same set of scores, the *z* score for a score of 48 would be

$$z = \frac{48 - 60}{8} = \frac{-12}{8} = -1.5.$$

Note that if the score is greater than the mean, the value of *z* will be positive. Conversely, if the score is less than the mean, *z* will be negative. If the score is equal to the mean, *z* will equal 0. In addition, the greater the difference between the score and the mean, the larger the absolute value of *z* will be.

Other standard scores can be derived from *z* scores by changing the value of the mean and standard deviation to some other desired values. This is done on most *norm-referenced standardized tests*, in which people's scores are interpreted in comparison to those of a *norming sample* of people believed to be representative of those who will take the test. *Transformed* or "dressed up" standard scores may be used because few people would want to hear that their (or their child's) score on an intelligence test or other standardized test was 0 or 1, let alone a negative number. For example, some intelligence tests use standard scores for which the mean has been transformed or reset to 100 and the standard deviation has been transformed to 15. Thus, a person whose score was equal to the mean of the norming sample would receive a score of 100. Another whose score was 1 standard deviation above the mean would receive a score of 115.

The equation for transforming a *z* score into another standard score (SS_t) with a mean of M_t and a standard deviation of SD_t is

$$SS_t = M_t + (z \times SD_t).$$

Thus, if a person got a score on an intelligence test that yielded a *z* score of −0.4 (i.e., 0.4 standard deviations below the mean), the transformed standard score would be

$$SS_t = M_t + (z \times SD_t) = 100 + (-0.4 \times 15)$$
$$= 100 + (-6) = 100 - 6 = 94.$$

Note that any pair of values can be used to designate the mean and standard deviation, and many have been. For example, *T scores* have a mean of 50 and a standard deviation of 10. Thus, a *z* score of –0.4 would yield a *T* score of 46 because $50 + (-0.4 \times 10) = 50 - 4 = 46$. The important thing to remember, however, is that no matter what form standard scores take, they always communicate the same two pieces of information:

1. Whether the score is above or below the mean, and

2. How far the score is from the mean.

Standard scores have many applications besides standardized tests, but there is an important requirement for their use: Standard scores can be used only

when the differences between scores are meaningful and consistent across the entire scale. It is important to realize that *measurement* (i.e., the characterization of a property or characteristic of a person, event, or thing as a numerical value) takes many different forms. The concept of *scales* (or *levels*) *of measurement* captures some critical differences. There are four scales of measurement: *nominal, ordinal, interval*, and *ratio*, and of these, only the latter two are appropriate for use in calculating standard scores. This is because it is only for scores with interval and ratio scales that differences between scores are meaningful and consistent as required for the appropriate use of the mathematical operations (e.g., addition, subtraction, multiplication, and division) used in finding the mean and standard deviation. Thus, you couldn't use standard scores to describe people's religious preferences (nominal) or the hardness of minerals (ordinal). You could, however, use standard scores to characterize temperatures whether they were measured on a Fahrenheit thermometer (an interval scale with no absolute zero, because a temperature of 0° Fahrenheit does not indicate the absence of heat) or a Kelvin thermometer (a ratio scale, i.e., an interval scale with an absolute zero, because a temperature of 0° Kelvin indicates the absence of heat). It should be noted that the *percentile ranks* (a measure of the proportion of scores lower than a given score) are often reported along with standard scores. Percentile ranks are measured on an ordinal scale and do not have the equal interval property of standard scores. *Stanine scores*, which assign all scores a value of 1–9, with 1 designating the lowest scores, are another form of ordinal measurement.

Ernest T. Goetz

See also Descriptive Statistics; Measurement; Norm-Referenced Tests; Percentile Rank; Stanine Scores; *T* Scores

Further Readings

Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice Hall.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory.* New York: Harcourt Brace Jovanovich.

Lyman, H. B. (1998). *Test scores and what they mean* (6th ed.). Boston: Allyn & Bacon.

Shavelson, R. J. (1996). *Statistical reasoning for the behavioral sciences* (3rd ed.). Boston: Allyn & Bacon.

STANFORD–BINET TEST

The Stanford–Binet test is arguably the premier test of intelligence. The latest version of the test is the Stanford–Binet Intelligence Test, Fifth Edition (SB5). The original Stanford–Binet test resulted from the efforts of Lewis Terman at Stanford University. He based the test on the early intelligence tests of Alfred Binet and Theophile Simon from France near the advent of the 20th century.

Binet and Simon constructed the first intelligence test in 1905 with subsequent revisions in 1908 and 1911. Educational researchers and psychologists in America subsequently became aware of the work of Binet and Simon. One such researcher was Terman, who became interested in composing a version of the test written in English and adapted to American individuals. In 1916, he prepared the first version of the test to assess intelligence for a wide range of individuals from young childhood to young adulthood. A related use of the test was the identification of individuals with limited intellectual abilities. Test administrators used the test to assess many European immigrants in the first two decades of the 20th century. Although those assessments were often invalid because many of the immigrants could neither read nor write English text, they still provided evidence for some of the inherent racial and cultural inferiority of certain immigrant groups such as emigrants from the Mediterranean region.

After the publication of the first edition of the test, Terman continued work on the test in an attempt to improve the predictive qualities of the test. Working in collaboration with Maud Merrill, Terman published a newer version of the Stanford–Binet test in 1937. The test had two parallel forms: one named Form L (for Lewis) and the other named Form M (for Maud). The correlation between total scores of the two forms was approximately .90, which attested to the reliability of the test. Merrill continued work on the Stanford–Binet test and constructed a new single form termed Form L-M that made use of items from Forms L and M. Her work and that of others led to the subsequent editions: the third edition in 1960, the fourth edition in 1986, and the fifth edition (SB5) in 2003.

There is an informative exposition of the SB5 on the Web site of the publisher, Thomson Nelson. There are 10 primary subtests in the SB5 that assess five factors of cognitive ability and two domains of

intelligence. The five factors of cognitive ability are fluid reasoning, knowledge, quantitative reasoning, visual-spatial reasoning, and working memory. Fluid reasoning relates to the capacity to abstract patterns. Knowledge relates to conceptual information. Quantitative reasoning relates to competency with arithmetic. Visual-spatial reasoning relates to facility with spatial transformations and visual forms. Working memory relates to the amount of new information that one can recall at any one time. The two domains of intelligence are nonverbal intelligence and verbal intelligence.

SB5 has a balanced, relatively simple structure relating subtests to domains of intelligence and factors of cognitive ability. Five of the 10 subtests assess nonverbal intelligence, and the other five assess verbal intelligence. For each of the five factors of cognitive ability, one can find 2 of the 10 subtests to assess that factor: a nonverbal subtest and a verbal subtest. For example, the two measures of fluid reasoning are the Nonverbal Fluid Reasoning subtest and the Verbal Fluid Reasoning subtest. The two measures of knowledge are the Nonverbal Knowledge subtest and the Verbal Knowledge subtest.

For each of the subtests, there are items that present activities for the testee. For example, to assess nonverbal fluid reasoning, there are items that assess the extent to which testees are able to identify objects that complete series of objects or matrices of objects. To assess verbal knowledge, vocabulary items measure the word knowledge of a testee. To assess verbal working memory, items assess the extent to which testees are able to remember whole sentences.

A test scorer can score the item responses by hand or with the use of a computer. The test scorer combines the subtest scores to form various composite scores, such as a factor index score, a domain score, and the Full Scale score. Each composite score is an IQ score with a mean of 100, a standard deviation of 15, and a possible range of 40–160. Thus, an individual may have a Nonverbal Intelligence IQ of 120, a Verbal Intelligence IQ of 90, a Fluid Intelligence score of 120, a Knowledge score of 100, a Quantitative Reasoning score of 95, a Visual-Spatial Reasoning score of 100, a Working Memory score of 90, and a Full Scale Intelligence IQ of 105. Such scores would indicate that the individual has above-average nonverbal and fluid reasoning skills and below-average verbal reasoning skills and working memory capacities.

The range of potential testees with the SB5 is vast—from 2 years to more than 80 years of age. The SB5 is an individually administered measure of intelligence and basic cognitive abilities that requires a trained tester for proper administration. The administration time is approximately 5 minutes per subtest.

The SB5 is useful in assessing a wide range of individuals at practically any age. The test items that measure nonverbal intelligence are informative in assessing subjects with hearing deficits and limited English proficiency. A tester identifies individuals with learning disabilities when there are substantial discrepancies between the nonverbal and verbal ability scores of the individuals.

The many challenging items are useful in identifying gifted and talented subjects. The many items of low difficulty are useful in identifying low-functioning adolescents and adults with mental limitations. Both educators interested in regular education and educators interested in special education and gifted education will find the SB5 to be an informative, diagnostic measure of intelligence and cognitive abilities. As noted by Bain and Allin, the SB5 is an exceptional measure of intelligence that is objective, reliable, and valid and that has a wide range of uses with a wide range of individuals.

William M. Bart and David P. Peterson

See also Cognitive View of Learning; Intelligence and Intellectual Development; Intelligence Quotient (IQ); Intelligence Tests

Further Readings

Bain, S. K., & Allin, J. D. (2005). Book review: Stanford–Binet Intelligence Scales, fifth edition. *Journal of Psychoeducational Assessment, 23,* 87–95.

Flanagan, D., & Harrison, P. (2005). *Contemporary intellectual assessment: Theories, tests, and issues* (2nd ed.). New York: Guilford.

Roid, G. H. (2003a). SB5 Scoring Pro [Computer software]. *Stanford–Binet Intelligence Scales, fifth edition.* Itasca, IL: Riverside.

Roid, G. H. (2003b). *Stanford–Binet Intelligence Scales, fifth edition, examiner's manual.* Itasca, IL: Riverside.

Roid, G. H. (2003c). *Stanford–Binet Intelligence Scales, fifth edition, technical manual.* Itasca, IL: Riverside.

Thomson Nelson. (n.d.). *Stanford–Binet Intelligence Scale, fifth edition.* Retrieved February 27, 2007, from http://www.assess.nelson.com/test-ind/stan-b5.html

STANINE SCORES

Stanine scores are area-normalized standard scores that are standardized to have a mean of approximately 5 and a standard deviation of approximately 2. They are simple to define but may discard some of the detail in the original data.

Many statistical procedures assume that the data are normally distributed. Sometimes, nonlinear algebraic transformations such as logistic or arcsin transformations will produce the desired result, but when the non-normality is due to sampling or to the features of the measuring instrument itself, an area-normalizing transformation may be required. The stanine transformation provides a quick and simple way to accomplish this goal.

As with other area normalizations, the stanine transformation *assigns* transformed scores to be equivalent to the raw scores rather than computing the new scores. To obtain stanine scores, carry out the following steps:

1. Make a frequency distribution of the raw scores.

2. Assign stanine scores as follows: Assign a value of
   - 1 to the lowest 4% of scores,
   - 2 to the next 7% of scores,
   - 3 to the next 12% of scores,
   - 4 to the next 17% of scores,
   - 5 to the next 20% of scores,
   - 6 to the next 17% of scores,
   - 7 to the next 12% of scores,
   - 8 to the next 7% of scores,
   - 9 to the top 4% of scores.

Stanine scores have two main advantages over other area-normalizing transformations. First, they are easy to compute and do not require a table of the normal distribution z scores. Second, because single scores cover a relatively wide portion of the score distribution, stanine scores tend to discourage users from overinterpreting relatively small differences in scores. This is one reason many test publishers offer stanines as one metric in which test results are provided.

A third advantage, which is no longer very relevant, is that stanines require only one column. At the time they were invented, when punch cards were the primary means of data storage and sorting and hand computation was required, single-digit data values offered a major saving in space and computational labor. The primary disadvantage of stanines is that they may discard some useful detail in the data.

Robert M. Thorndike

See also Descriptive Statistics; Standard Deviation and Variance; *T* Scores

STATISTICAL SIGNIFICANCE

Perhaps the most recognized type of significance testing is statistical significance. The concept of significance dates back to 1710, when John Arbuthnot, an English physician, published his statistical analysis. Statistical significance tests point to the probability of obtained sample results that deviate from the population specified by the null hypothesis, in a particular sample size. A *null hypothesis* is the hypothesis that is to be tested. It is important to clarify that statistical significance tests do not evaluate the probability that the sample results represent the population. In actuality, statistical significance tests work on the assumption that the null hypothesis describes the population and then test the sample's probability.

Probability

Statistical significance testing may be used when working with a random sample from a population, or a sample that is believed to approximate a random, representative sample. Statistical significance testing calls for subjective judgment in establishing a predetermined probability (ranges between *0* and *1.0*) of making an inferential error caused by the sampling error. When using statistical significance testing, it requires the use of two forms of probability (P): calculated and critical. Statistical significance is met when $P_{(CALCULATED)}$ is less than $P_{(CRITICAL)}$; when this is the case, the null hypothesis can be "rejected." Only when the null hypothesis is rejected are the results called "statistically significant." This simply implies that the sample results are relatively unlikely, given the assumption that the null hypothesis is exactly true.

$P_{(CRITICAL)}$

One of the probabilities, $P_{(CRITICAL)}$, is also referred to as "alpha." $P_{(CRITICAL)}$ is the probability of making a Type I error when testing a null hypothesis. A Type

I error occurs when rejecting a null hypothesis that is true. Another possible type of error that can occur is a Type II error, which occurs when the null hypothesis is not rejected and it is false. The $P_{(CRITICAL)}$ is usually set before collecting the data and tends to be a small number. The most frequent alpha levels used are .05 and .01; by using a small number, the probability of error is minimized.

$P_{(CALCULATED)}$

The other form of probability is $P_{(CALCULATED)}$, which also ranges between *0* and 1.0. Probabilities can be calculated only in the context of assumptions sufficient to constrain the computations such that a given problem has only one answer. While it is easy to set a value for $P_{(CRITICAL)}$, calculating $P_{(CALCULATED)}$ can be difficult. Therefore, test statistics (e.g., F, t, X^2) have been used instead because they are easier to calculate and more convenient reexpressions of $P_{(CALCULATED)}$.

Misinterpretation of Statistical Significance

When working with statistical significance tests, it is important to be aware of three common misinterpretations. The first misinterpretation is that $P_{(CALCULATED)}$ is an indication of how likely the sample results describe the population. $P_{(CALCULATED)}$ does not inform the replicability of the results. The second misinterpretation is that $P_{(CALCULATED)}$ is the probability that the results were due to chance, and that smaller *p* values indicate stronger evidence that the null hypothesis is false. This is a misinterpretation because it has been demonstrated that $P_{(CALCULATED)}$ can be smaller than the probability of the null hypothesis being true. The third misinterpretation is that $P_{(CALCULATED)}$ is the probability that the null hypothesis is true. Despite the common misinterpretations, statistical significance testing continues to be widely used even when it does not tell us what we want to know. Scholars suggest that users of statistical significance tests do not fully understand what these tests do, and test users mistakenly believe that statistical significance tests evaluate whether results were due to chance.

Criticism of Statistical Significance

Other criticism of statistical significance tests is its relationship with sample size. With a large enough sample size, statistical significance can always be achieved.

Taking this into account, it is important to recognize that statistical significance does not equate to practical significance, clinical significance, or importance. Additionally, researchers should not rely on statistical significance testing to determine the value or importance of results because it is possible for results not to reach statistical significance, yet be of importance and replicable. In some instances, data analyses that do not meet the .05 level of significance could make a meaningful contribution to the literature but may not be submitted for publication (i.e., file drawer problem). Therefore, it is important to remember that although results may not reach the commonly used .05 level of statistical significance, they may still hold value.

When analyzing psychotherapy or medical outcomes, statistical significance has little clinical relevance. The following is an example that demonstrates how statistical significance does not produce clinical relevance or importance. In a study for a fever-reducing drug, patients who had an average temperature of 104° F were placed into a treatment and a control group. Results of the study found that the fever-reducing drug produced a statistically significant temperature reduction, with an average temperature reduction of 0.7° F. Although the findings were statistically significant, the fever-reducing drug did not result in improvement for the patients because they still had an average temperature of 103.3° F. Clearly, it was not enough for the outcome to be statistically significant. In this situation, other significance tests or other ways of determining significance are needed to determine the degree of impact the treatment has had on the patient or client.

Alternatives to the Statistical Significance Test

Practical Significance

Unlike statistical significance, practical significance determines if results are useful in the real world. Results are determined to be useful by looking at *effect magnitude*, which comprises measures of strength association and effect size.

Clinical Significance

Clinical significance is the practical or applied value or importance of the effect of the intervention.

Miguel Ángel Cano

See also Inferential Statistics; Standard Scores; *T* Scores

Further Readings

Chow, S. L. (1996). *Statistical significance: Rationale, validity and utility.* Thousand Oaks, CA: Sage.

Thompson, B. (2002). "Statistical," "practical," and "clinical": How many kinds of significance do counselors need to consider? *Journal of Counseling and Development*, *80*, 64–71.

Thompson, B. (2006). *Foundations of behavioral statistics: An insight-based approach.* New York: Guilford.

STEREOTYPES

The term *stereotype* was first used in the printing industry, where it referred to a cast iron plate used to make repeated impressions of the same image. This sense of an impression that is enduring, fixed, and resistant to change was imported into the social sciences by the journalist Walter Lippmann in his book *Public Opinion.* This applied the term to refer to the "pictures in the head" that people have of social groups.

A stereotype is an image or representation of a group of people that is widely known and shared within a particular community or group. However, in addition to this, within social research (as in society more generally), stereotypes are typically understood to have a number of negative features, many of which were first identified in Lippmann's writing. In particular, they are seen as akin to caricatures, which, as well as being resistant to change, are also biased, selective, and simplistic.

Consistent with these views, a large body of social psychological research confirms that stereotypes tend to accentuate both the differences between groups and the similarities within them. Their content is also often observed to be ethnocentric in the sense that the stereotypes of groups to which one belongs (ingroups) tend to be more positive than those of groups to which one does not belong (outgroups). For this reason, when most social scientists use the term *stereotype*, they are implicitly referring to perceptions that are believed to be fundamentally inaccurate—representing members of outgroups as being more similar to each other, more different from other groups, and of worse character than they really are: as if "we are all good" and "they are all bad." For this reason, holding and expressing stereotypes is often seen as a form of prejudice, and this has contributed to an intertwining of research into these two phenomena.

The history of research into stereotypes and stereotyping has gone through a number of clearly defined phases. As a result, research into this topic has been an important context for key debates in social (and educational) psychology and has been a site for important conceptual and theoretical advances. The earliest research, pioneered by Daniel Katz and Kenneth Braly in the early 1930s, examined stereotype consensus and content by asking people to assign adjectives from a long list to members of a range of national and ethnic groups, including their own. This confirmed that some stereotypes were indeed widely shared and that some were very disparaging. Thus, whereas Americans were most likely to describe Americans as industrious and intelligent, they described "Jews" as shrewd and mercenary and "Negroes" as superstitious and lazy.

However, against the view that stereotypes are inherently fixed and prejudicial, later studies showed that stereotype content changed to reflect changes in the nature of intergroup relations. During World War II, for example, Americans' stereotypes of Germans and Japanese became much more negative, but these improved once the conflict was over. Likewise, in Muzafer Sherif's famous studies of boys attending a summer camp, stereotypes of different groups changed to reflect the nature of the relations between them. So, when the groups were in conflict, stereotypes of the outgroup were much more negative than they were when the groups needed to cooperate to achieve a superordinate goal. Later studies also showed that stereotypes of one group changed substantially depending on with which other groups it was compared.

In the aftermath of World War II, a large amount of energy was focused on the question of whether people with particular personalities were more likely to exhibit black-and-white thinking of the type associated with stereotypes. This work was given impetus by the research of Theodor Adorno and his colleagues into the "authoritarian personality," which was particularly concerned with explaining the development of anti-Semitic stereotypes in 1930s Germany. A key part of these researchers' answer was that stereotypic thinking was the hallmark of people who had a propensity for black-and-white thinking and who were intolerant of ambiguity.

Again, though, the evidence did not support this view or personality-based models of stereotyping more generally. So, although initially popular (and fitting with lay views about the pathological nature of

those who are prejudiced), research provided little support for the argument that only authoritarians endorse stereotypes. It is certainly the case that people who do not have right-wing leanings (and hence are unlikely to embrace an authoritarian ideology that favors stratification and deference to those in authority) hold different stereotypes from authoritarians, but they stereotype nonetheless. Moreover, explanations that are couched in terms of individuals' personalities fail to explain the collective dimensions of the stereotyping process. The development of anti-Semitic stereotypes in Nazi Germany was socially and politically significant not because these views were held by a select few, but because they were embraced by a large populace. Indeed, it was this sharedness that made the stereotypes so potent.

These realizations led researchers to move away from an analysis that saw stereotyping as a process specific to particular individuals. A particularly influential figure in this movement was Gordon Allport, who followed Walter Lippmann in observing that members of all groups hold and use stereotypes. Moreover, he argued that stereotyping was a normal cognitive activity that was essential for a predictable and manageable life and that derived from the rational (if error-filled) process of categorization. These insights were distilled into the view that stereotypes are a form of "necessary evil": They are the outcome of a simplification process that arises from the cognitive impossibility of treating everyone as an individual, but that, as a result, also introduces distortion and bias.

This latter view was given further impetus by the empirical work of Henri Tajfel in the 1960s. This provided evidence that cues that encourage people to perceive objects as members of distinct categories lead to the accentuation of intercategory difference and intraclass similarity. So, in a series of lines of different lengths, if the longer lines are labeled "A" and the shorter lines are labeled "B," then there is a tendency to accentuate the difference in the length of the two sets and the similarity within them. In this way, once viewed as part of a common category, different stimuli (whether lines or people) tend to be seen both as more similar to each other and as more different from other stimuli than they would be if the category had not been invoked.

The insights that this analysis provided fueled a third wave of social cognition research in the 1980s. An idea that came to be central to work in this tradition was that people stereotype because they are "cognitive misers" whose perceptions and judgments are driven by the all-important goal of saving precious information-processing resources. Led by researchers such as Susan Fiske and David Hamilton, work in this tradition sought to identify a range of normal cognitive processes that make stereotypes efficient but full of errors.

In this vein, an abundance of experimental studies suggested that stereotypes arise automatically from a range of energy-saving, information-processing biases that come into play at every stage of the stereotyping process—from encoding and storage to retrieval and communication. Researchers argued that these biases lead people to (a) make faulty associations between attributes and social groupings (so-called illusory correlations); (b) attend to novel and distinctive information, as well as to that which is consistent with their expectations and prior theories (the self-fulfilling prophecy); and (c) emphasize the homogeneity of outgroups and the heterogeneity of ingroups (the outgroup homogeneity effect). The most recent cognitive work has also sought to explore the differences between aspects of these processes that are automatic and those that are under the perceiver's conscious control.

Although the social cognitive approach to stereotypes and stereotyping remains highly influential, it has also been critiqued by researchers who argue that such research neglects the role that social context plays in shaping and structuring cognition. In particular, social identity and self-categorization theorists such as Penelope Oakes, S. Alexander Haslam, and John Turner have argued that stereotypes are primarily political tools that allow groups to represent and influence the reality of intergroup relations as perceived from their particular social vantage point. According to this view, stereotypes exist not to save effort but to make social and political behavior possible. Hence, where it is observed, the error of stereotypes is seen to have its basis not in psychological deficiency but in the political positions and aspirations of the groups that hold them. Moving away from the view that stereotypes are the product of more-or-less automatic, intrapsychic processes, this research has led to renewed interest in group processes (e.g., of communication and social influence) that contribute to stereotype consensus and help bring about stereotype change.

This plurality of perspectives has ensured that stereotyping research has remained at the heart of debate in social psychology and that it continues to drive both theoretical and empirical advance. The centrality of stereotypes to an array of social, political,

and educational issues also ensures that there continues to be a keen interest in these developments from researchers and practitioners outside social psychology.

S. Alexander Haslam

See also Discrimination

Further Readings

Allport, G. W. (1954). *The nature of prejudice.* Cambridge, MA: Addison Wesley.

Fiske, S. T., & Taylor, S. E. (1984). *Social cognition.* Reading, MA: Addison Wesley.

Lippmann, W. (1922). *Public opinion.* New York: Harcourt Brace.

McGarty, C., Yzerbyt, V. Y., & Spears, R. (Eds.). (2002). *Stereotypes as explanations: The formation of meaningful beliefs about social groups.* Cambridge, UK: Cambridge University Press.

Oakes, P. J., Haslam, S. A., & Turner, J. C. (1994). *Stereotyping and social reality.* Oxford, UK: Blackwell.

Schneider, D. J. (2004). *The psychology of stereotyping.* New York: Guilford.

Spears, R., Oakes, P. J., Ellemers, N., & Haslam, S. A. (Eds.). (1997). *The social psychology of stereotyping and group life.* Oxford, UK: Blackwell.

Tajfel, H. (1969). Cognitive aspects of prejudice. *Journal of Social Issues, 25,* 79–97.

Tajfel, H. (1981). Social stereotypes and social groups. In J. C. Turner & H. Giles (Eds.), *Intergroup behaviour* (pp. 144–167). Oxford, UK: Blackwell.

STIMULUS CONTROL

A primary objective of psychology is the prediction of behavior. Operationally, this means that what an individual will do in a particular situation can be anticipated. To the degree that this is possible, the individual is under stimulus control. The fact that people seldom are surprised by the actions of those with whom they interact is an indication that others are behaving as expected. This entry briefly discusses how this stimulus control is acquired. It progresses from simple reflexive actions to those resulting from the differential consequences of one's behavior in different situations. For example, one learns to use different language and manners when speaking to a teacher than when speaking to a fellow student.

The simplest form of stimulus control is those unlearned responses naturally elicited by an unconditioned stimulus (UCS). Thus, an air puff to the eye will produce an eye blink, contact with a hot stove will result in rapid retraction of one's hand, and food in the mouth will elicit salivation. However, as I. P. Pavlov showed, this control can be transferred to stimuli that precede, and therefore predict, these naturally eliciting events. For example, stimulus control is revealed when a dog salivates to a tone that precedes food and when the child who burned his or her hand on the hot stove becomes afraid of stoves. Both are examples of classical conditioning and illustrate that this type of associative learning is capable of (a) bringing a particular reflexive response under the control of a conditioned stimulus (CS) that was originally neutral with respect to that behavior, and (b) creating motive states related to the signaled events—repulsion ("fear") in the case of aversive events and attraction with those that are positive. Recall being drawn into a bakery when you had no interest in food prior to experiencing the aroma of your favorite, fresh-baked pastry. This *incentive-motivation* is clearly under the control of stimuli signaling positive (attractive) or aversive (repelling) events significant to the individual.

Traditionally, students of stimulus control have been most interested in operant conditioning that is concerned with when, whether, and in what way an individual's behavior will be affected by particular environmental consequences or events related to that behavior. C. B. Ferster and B. F. Skinner pioneered the study of these relationships, whereby operant reinforcement contingencies describe the relationships between emitted behavior and consequences. Those operant behaviors that produce a reinforcer (e.g., an attractive event such as food, approval, or a safety signal) will increase in frequency. In contrast, behaviors that produce an aversive situation (e.g., a repelling event such as shock, rejection, or a time-out from reinforcement signal) decrease the probability of those behaviors.

In nature, there tend to be cues in the environment that signal when a specific contingency is in operation and one's behavior will therefore produce the consequence(s) specified by that contingency. Differential reinforcement of a particular behavior pattern under a limited range of environmental conditions is what produces operant *stimulus control.* For example, pressing a lever may produce food pellets for a deprived rat only when a house light is on, just as a speaker is likely

to get an answer to his or her question only when he or she has a listener's attention. This can be represented schematically by the three-term contingency: SD • R → Sr, which states that a behavior (R) (e.g., asking a question), emitted in the presence of a discriminative stimulus (SD) (e.g., an attentive, prepared listener), will produce a reinforcer (Sr) (e.g., an answer). This also implies that when the SD is absent (e.g., a listener isn't present, or the one who is isn't attentive or prepared), these responses will be ineffective.

By now, it is probably becoming clear to the reader that a major objective of the educational process is bringing students' verbal repertoire under stimulus control so that they can give appropriate answers to the questions and intellectual challenges posed to them. Discrimination learning is central to this process, and that requires that the teacher be able to differentially reinforce such learning. Attention, praise, and grades are major potential reinforcers at the teacher's disposal. However, these are "arbitrary" reinforcers in that they are not intrinsically related to the goals of education. Thus, a primary objective of the educational process is to make learning "generically" reinforcing—that is, reinforcing in its own right. For different individuals, playing basketball, doing research, creating art, skiing, and/or reading in an area of one's interest could be examples of behaviors that are generically reinforcing. A teacher might endeavor to make learning in general satisfying by systematically expanding incrementally upon interests with which a student enters the class. In addition, he or she might apply the Premack Principle, whereby less preferred behaviors are reinforced by creating the opportunity to engage in more preferred behaviors.

Unfortunately, being able to predict when a particular operant behavior will be emitted does not necessarily reflect a complete appreciation of the multiple factors contributing to the stimulus control of behavior in any particular situation. The author demonstrated this, while illustrating the potential subtlety of operant contingencies, by training two groups of rats to press a lever when either a tone or a light was present and not press the lever when these stimuli were absent ($\overline{T}+\overline{L}$). They differed in that (a) one group's responses during the tone or light produced food while food was unavailable in $\overline{T}+\overline{L}$, whereas (b) the other group had to respond during the tone or light to enter $\overline{T}+\overline{L}$, where they received food for not responding. These groups were behaviorally indistinguishable, responding in tone or light but not in $\overline{T}+\overline{L}$. But

were they under comparable stimulus control? A stimulus compounding assay revealed that they indeed were not. Presenting the tone and light simultaneously (T + L) tripled the response rates of the rats who had received food during tone and during light, whereas it had no effect on the response rates of the rats who had never received food during these stimuli. Why?

The behavior occasioned by an SD primarily reflects the accompanying operant contingency effective therein. That contingency defines what behavior(s) can ultimately produce reinforcement. This represents the response-discriminative process. However, an SD also signals reinforcement changes that will activate the classically conditioned incentive-motive process. An SD associated with reinforcement increase will produce an excitatory, energizing, incentive-motive state—like the bakery aroma mentioned earlier. One associated with a decrease in reinforcement will produce an inhibitory, suppressing incentive-motive state. In the experiment described above, the stimulus compounding test was necessary to reveal the operation of the response-discriminative and incentive-motive processes. When both were increasing, T + L tripled responding, but when they were in conflict (where the tone and the light occasioned an increase in response while signaling a decrease in reinforcement), T + L didn't change the response rate compared to that controlled by tone or light alone. Both processes are potentially activated in all multicomponent operant schedules of reinforcement. Therefore, in nature, it is likely that most classically derived incentive-motivation is created while the organism is under discriminative stimulus control on operant baselines rather than through CS–UCS pairings that occur independently of behavior.

Research has also revealed that contingencies experienced in the past (i.e., conditioning history) can influence current stimulus control even when baseline behavior is appropriate to the contingencies currently operating. Stanley Jerome Weiss demonstrated this in an experiment where two groups of rats appeared behaviorally comparable, responding to avoid a shock in tone and not responding in light that was shock free. However, in only one group did the safety signal (light) appear to inhibit "fear" when presented with the tone, reducing the avoidance rate by 50%. For this group, where light became a fear inhibitor, the light had always been safe. In the other group, light had a shock-related history before it was made a safety signal—and that history eliminated its capacity to reduce the

tone-produced fear even though the light no longer occasioned avoidance. This suggests that a stimulus compounding assay of a fear reduction treatment's effectiveness should be employed, in addition to just the simple elimination of avoidance, as a test of treatment effectiveness. This again illustrates why actions are often misinterpreted, and how behavior patterns that appear comparable could have resulted from different combinations of underlying learning-derived processes. Appreciating that complexity is one of the significant challenges in creating effective education.

Stanley Jerome Weiss

See also Classical Conditioning; Operant Conditioning

Further Readings

Dinsmoor, J. A. (1995). Stimulus control: Part I. *Behavior Analyst*, *18*, 51–68.

Dinsmoor, J. A. (1995). Stimulus control: Part II. *Behavior Analyst*, *18*, 253–269.

Ferster, C. F., & Culbertson, S. (1982). *Behavior principles* (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.

Johnson, K. R., & Layng, T. V. J. (1992). Breaking the structuralist barrier: Literacy and numeracy with fluency. *American Psychologist*, *47*, 1475–1490.

Morris, E. K., & Redd, W. H. (1975). Children's performance and social preference for positive, negative, and mixed adult-child interactions. *Child Development*, *46*, 525–531.

Sidman, M. (1986). Functional analysis of emergent verbal classes. In T. Thompson & M. D. Zeiler (Eds.), *Analysis and integration of behavioral units* (pp. 213–245). Hillsdale, NJ: Lawrence Erlbaum.

Weiss, S. J. (1978). Discriminated response and incentive processes in operant conditioning: A two-factor model of stimulus control. *Journal of the Experimental Analysis of Behavior*, *30*, 361–381.

Weiss, S. J., & Schindler, C. W. (1985). Conditioning history and inhibitory instrumental stimulus control: Independent-groups and within-subjects measures. *Animal Learning & Behavior*, *13*, 215–222.

Weiss, S. J., & Schindler, C. W. (1987). The composite-stimulus analysis and the quantal nature of stimulus control: Response and incentive factors. *Psychological Record*, *37*, 177–191.

Students' Rights

The general concept of students' rights is that every child has the right to receive a free and appropriate education, and through this educational process, students have rights in many areas, including expression, discipline, safety, privacy, treatment, and accommodations. Therefore, students' rights are an important topic in any discipline that deals with education.

The focal point of all educational activity is the students. Teachers, principals, school social workers, school psychologists, and other staff are all put in place to influence the process by which students receive their education. Regardless of age, disability, race, sexual orientation, and so on, every student has the right to be treated and encouraged with the idea that he or she has value and worth, and the education should be provided as a component of anticipated success for the student.

Many professionals feel that students have too many rights and that those rights hinder the educational process. Some teachers and schools fear lawsuits and other legal trouble. Schools can be sued over questionable discipline practices. This may result in limiting disciplinary procedures and giving a small number of disruptive students the potential to control the classroom environment. In some cases, students have been known to report teachers by making false accusations. Any accusations made, true or fabricated, have the possibility of ending up in a lengthy and expensive court process. On a positive note, these restrictions help protect students and also welcome more creative positive reinforcement techniques.

Along with the diversity of the student bodies across the country comes a broad spectrum of circumstances that put stipulations on how the basic right of a free and appropriate education is materialized. Certain factors such as a student's neighborhood, race or ethnicity, socioeconomic status, level of disability, and safety put limitations and restrictions on the way that students' education is provided and their rights ensured. Local funding causes the quality of education to vary according to the district in which the student lives. A student's educational experience in a wealthy suburb may look very different from a student's experience in a poor urban district. In addition, a student's race, ethnicity, or socioeconomic status may lower the expectations a teacher of another background places on that child. Again, this jeopardizes a student's access to fair educational practices.

There are various areas of student rights, and they receive attention to different degrees. There is a common theme of balancing rights on one side with different rights on another side. If a student's individual

rights are protected, will that interfere with other students' rights? Through the blurred boundaries emerge some general rights that students receive, such as those pertaining to expression, discipline, suspensions and expulsions, corporal punishment, searches and seizures, school violence and safety, confidentiality, discrimination, and special needs.

Freedom of Expression

A student's right to free expression is one of the most frequently occurring topics in the literature. These rights include many areas such as dress code, school publications, speeches, conversations, and even Internet communication. With so many different areas of expression, it is often challenging to have a clear standard of policies and procedures as it pertains to a student's use of communication. In addition, laws governing expression are flexible. This flexibility, however, has the potential to strictly limit a student's desire to express himself or herself freely and to develop as an individual through this expression.

There is a constant requirement for schools to balance the need for a safe school environment and the right of students' expression. It is important to respect the rights of students, but on the other hand, these students are under the protection and care of adults who work in the schools.

The First Amendment of the U.S. Constitution guarantees freedom of speech; however, cases have been made that students do not necessarily enjoy this privilege to its full extent. The Supreme Court has gradually limited students' rights in terms of speech. However, if a student's speech is not a true threat, authorities are obligated to allow the student to express him- or herself freely. Students' freedom of expression cannot be upheld if such expression threatens the general operation of the school or leads to the denial of other students' rights. In addition, schools have the right to eliminate speech that is inappropriate for the student audience. A student's speech must be consistent with the values of the school, and the school officials have the leverage to decide which speech is inappropriate and unacceptable. School staff members are given the power to decide if a student's speech will jeopardize the accepted values and practices in a particular school setting.

Schools also have the power to regulate speech that may be representative of the school even if it does not happen on the school grounds. If a student's expression during out-of-class times has the potential to disrupt the educational process or endanger other students, school authorities have the power to limit that expression. One area that schools have monitored is school newspapers or other publications that may contain material that is not suitable for the maturity level of the school audience. In this situation, school officials may prevent students from publishing whatever they feel, requiring that the questionable information be removed from the school publication. There have also been cases where students have had their free expression through the Internet regulated because of inappropriate communication about things such as school violence and hatred toward the school and specific staff.

Another area of expression that is regulated in some schools is dress. Certain school officials feel that loosely regulated dress codes hinder school progress. Some staff members feel that it can be a distraction when students are competing with each other on the latest fashion. Other students may come to school in inappropriate clothing that infers things of a sexual, prejudicial, or violent nature. These types of clothing expressions follow the same guidelines when it comes to students' rights. Students are allowed to use creative expression until it jeopardizes the rights of other students or disrupts the educational process of the school. These guidelines are based on the subjective decisions of the school officials who regulate this expression.

One of the major elements guiding the regulations on expression is the threat of violence. School staff members evaluate forms of expression for their potential as threats to the school atmosphere. If a student's expression is considered to interfere with the safety of the school and puts other students at risk, then school officials are allowed and obligated to stop this form of expression. School authorities have considerable leniency when it comes to these regulations.

Students in general seem to have a certain amount of freedom when it comes to their expression; however, their rights are limited when it comes to threatening or inappropriate expression. One benefit to this regulation is that students learn firsthand the complexity of society. They are able to learn how to handle their expression with a sense of responsibility and self-regulation, keeping in mind the common good of the entire student body when they are expressing themselves. On the other hand, there is a possibility of students losing opportunities for creative expression or even for cries for help because of these regulations.

Discipline

In any disciplinary situation, the student does have rights. In any circumstance, especially one that ends up in court, the student has a right to an advocate who is knowledgeable about the student's rights and can ensure that the student's side will be heard and respected during a hearing. Every student has value and rights regardless of the behaviors and actions in which he or she may have taken part.

In addition, general education students also have a right to appropriate education, and if the behaviors of a particular student are disrupting that environment, it may be necessary to remove that disruptive student from the general education classroom. Some students have disabilities that cause behaviors that are out of their control, yet still a hindrance to the teaching of the classroom. In situations such as this, the rights of the general student body may take precedence over the rights of that individual child. The individual child may be relocated to another classroom or even another school if doing so satisfies the rights of an entire classroom or school.

Suspensions and Expulsions

Suspensions and expulsions have become a popular form of discipline in the schools. In these situations, students still have rights. A central right during any forced absence from school is that students are given minimal due process proceedings. The students are to be made aware of the charges brought against them. Once these charges are brought to their attention, the students also have a right to dispute them. These rights all must be honored before even a short suspension. For longer suspensions and expulsions, even more complicated procedures are necessary to allow for students to exercise their full rights. In addition, students may not be counted absent and penalized academically for missing class during any forced time away from school.

These rights also carry over to students with disabilities along with additional rights that must be granted to this specific population of children. A student with a disability must not be denied free and public education because of a disciplinary procedure. This denial would happen through changing a student's placement as set out in his or her individualized educational plan. It could also happen if the student were to miss an excessive amount of school because of suspensions. And if the student with a disability displays behavioral problems that are a manifestation of his or her disability, he or she cannot be suspended for those actions. However, if the behavioral problems are not a manifestation of the student's disability, he or she may be disciplined in the same manner that a typical student would be for suspensions up to 10 days. However, if the suspension is for a longer period, the district and the student's parents must agree on an alternative setting. These protections also exist for students who are not yet identified as having a disability, but may in fact have a disability and may be eligible for special education services.

Corporal Punishment

It is not unconstitutional to use corporal punishment in the schools. As with many areas of student rights, school officials are expected to use their own discretion when it comes to the use of corporal punishment. The punishment must be proportionate to the action that led to the discipline. Also, the punishment must not be guided by malice, and it must not inflict severe injury on the student. Court decisions have guided the idea that school staff are allowed to use corporal punishment when it is necessary to discipline a student and to allow for proper education to take place. Although corporal punishment is constitutionally allowed, most states have laws against the use of corporal punishments in the schools. It is meaningful to note that only at the local level do some students have the right to not be physically punished while attending schools.

Searches and Seizures

Another major area in the literature as it relates to students' rights is searches and seizures. Increased attention on school violence and drug use has caused more desire and need to keep a watchful eye on the activities of students. School officials have gained a significant amount of leniency when it comes to searching and confiscating things from students. The laws and regulations around these practices are very vague and leave a lot of the judgment up to the school performing the search. The U.S. Constitution has yet another amendment that would normally protect students against this type of procedure. The Fourth Amendment allows citizens to be free from unreasonable searches, but as with the other amendments, this one is adjusted in the schools. The need to maintain

safe schools outweighs the rights of students to be free from searches. However, this does not mean that students have no rights when it comes to searches and seizures. This is another area where a balance must be achieved between the greater good of the school and each student's individual rights.

When students are being searched, there are different rules that school personnel must follow. Law enforcement officials are required to establish probable cause and must secure a warrant before performing a search. Police are allowed into the schools to question or arrest students, but students do have the right to remain silent during any questioning. Police are also not allowed to remove students from school unless they are placed under arrest. School staff, however, do not need a warrant or probable cause to search a student. These staff members only need reasonable suspicion to conduct searches. The reasoning behind this is that a threat may need immediate attention, and school staff may put the school and other students in jeopardy if they have to wait to obtain a warrant or justify probable cause. Schools must constantly balance the risk of being too hands-off and being too excessive when it comes to searching students. Schools need to be careful to maintain the students' rights, while at the same time protecting the school against potential disturbances.

During searches, certain procedures and policies must be followed. If a school official has reasonable suspicion to search a student, he or she must have facts displaying the basis for the search. He or she also must uphold appropriate search strategies as they relate to the student's gender and maturity level. In addition, the search should not be more intrusive than necessary. School and law enforcement officials may search students without reasonable suspicion if they are given voluntary consents by the students. These consents must take into consideration the age, education, and mental capacity of the students. However, it is not required that students be advised of their right to refuse to give consent.

Some scenarios call for different procedures to be followed. For example, when a school district has an established history of problems, it is allowed more search opportunities. School district personnel may perform random and targeted searches through measures such as metal detectors and parking lot sweeps in order to prevent weapons and drugs from being brought into schools. In addition, lockers are considered part of the school facility, so often there is little privacy provided to students when it comes to searching their lockers. Also, athletes and members of other school organizations have a lesser degree of privacy and are subjected to random drug testing. These scenarios display the flexibility that authority figures have when trying to maintain a safe and drug-free school system while at the same time keeping the best interests of the students in mind and upholding the students' rights to the greatest degree possible.

School Violence and Safety

Safety concerns also place restrictions on students' rights. After school shootings aroused public concern, there has been an increase in security in the schools. Students have a right to privacy and freedom of choice, but the potential for school violence has outweighed these rights in many circumstances. Security officers have, in many instances, free rein to search and question students even in the absence of suspicious activity. Many times, the idea of a greater good—safety and security in schools for all children—outweighs a student's right to privacy.

Confidentiality

Students have had records kept on them ever since they entered school. There are many rights involved with having these records on file. First of all, students and parents should be educated annually on their rights under the Family Educational Rights and Privacy Act. They have the right to view their files at any time and have the right to challenge any information they believe is false. Also, before any personal or confidential information is released to an unauthorized person, written consent must be obtained from the parent or legal guardian. Also, information should not be shared with individuals who do not have a need to know or have no legitimate interest in the child. Finally, only objective information should be placed in the student's file. Information on minor infractions that may change or information guided by opinion should not be in the student's file.

Discrimination

Students have the legal and ethical right to not be discriminated against in the school setting. Regardless of race, gender, or sexual orientation, all students are entitled to equal access to education and the opportunity to

learn in a safe environment. Members of racial minority groups have the right to have the same level of expectations placed on them as White children. These expectations include teachers not assuming a student of a certain race will not perform at the level of other students, teachers encouraging growth and enrichment in all students equally, and students being subjected to the same disciplinary procedures regardless of race. Girls and boys also are entitled to be provided an equal experience. It is illegal to provide extracurricular activities unequally to boys and girls. Additionally, girls cannot be removed from school for being pregnant. Sexual minorities also deserve protection against unfair treatment and must be provided an opportunity to explore themselves without being harassed and looked down on. Another right associated with discrimination is for minorities to have equal access to literature, history, and experiences that reflect their own background, not just the background of majority groups.

Special Needs Students

Another area of huge concern is the area of disabilities. Students with disabilities have the right to the least restrictive learning environment. However, these rights are blurred in certain situations. If an inclusive learning environment as part of the mainstream educational process is not conducive to a child's education, the student has a right to special services and/or educational settings that are more appropriate to his or her needs. This right often comes in conflict with budget issues, but this right is to be guaranteed by the district regardless of its financial situation.

All students have the right to a free and appropriate education, and students with disabilities have the same rights and have the right to experience this in the same schools and programs as students without disabilities. Students with disabilities have the right to an individualized education program that is developed and implemented with the goal of the student receiving education in the least restrictive environment and with all the needed services and accommodations to accomplish this. This is to be done at the school's expense and at no cost to the parent. Also, students have the right to evaluations that check the progress of their educational plan, and the parents and student have a right to participate and be a part of this planning. If the parents or student are in disagreement with the school's evaluation, they have the right to get an outside evaluation to determine the student's

eligibility for services. In addition, when a student receiving special education services reaches the age of legal consent, he or she must be privileged with all legal rights of a person his or her age regardless of the disability, if competent to do so.

Also, students have the right to be treated as individuals in the school setting. If interventions can keep a student out of special education, then those accommodations need to be made. Also, if a student is being tested for special education, the test should be administered in the student's native language so that he or she has an equal chance to perform to his or her ability on the test. Students' behaviors that are a manifestation of their disability should not be held as grounds for suspension and expulsion, but instead the parents and school staff should work together to advocate for the best placement for the student, whether it be the traditional school or an alternative placement. Students should also have a level of choice when it comes to their schooling. Children all learn differently, and whether they struggle or are extremely gifted, they deserve to be challenged and to be given the opportunity for growth through educational experiences that they find rewarding.

Future Implications

Students have basic rights that are meant to protect their well-being and educational attainment. Along with these rights come infinite possibilities for complications and adjustments. Special circumstances may lead to rights being altered or disregarded with a greater objective in mind. Education is meant to aid students in their journey to a fulfilling future, and every student has a right to take part in this experience with as little restriction and as much benefit as possible.

Mead Goedert and
Debra M. Hernandez Jozefowicz-Simbeni

See also Conflict; Discipline; Parent–Teacher Conferences

Further Readings

Beger, R. R. (2003). The "worst of both worlds": School security and the disappearing Fourth Amendment rights of students. *Criminal Justice Review*, 28, 336–353.

Cassidy, W. (2005). From zero tolerance to a culture of care. *Education Canada*, 45, 40–42.

Constable, R., McDonald, S., & Flynn, J. P. (1999). *School social work: Practice, policy, and research perspectives* (4th ed.). Chicago: Lyceum.

Ehlenberger, K. R. (2002). The right to search students. *Educational Leadership*, 59, 31–35.

Essex, N. (2004). Confidentiality and student records: Ten ways to invite legal challenges. *Clearing House*, 77, 111–113.

Hils, L. (2001). "Zero tolerance" for free speech. *Journal of Law & Education*, 20, 365–373.

Imber, S. C., & Radcliff, D. (2003). Independent educational evaluations under IDEA '97: It's a testy matter. *Council for Exceptional Children*, 70, 27–44.

Lindsey, P., Wehmeyer, M. L., Guy, B., & Martin, J. (2001). Age of majority and mental retardation: A position statement of the division on mental retardation and developmental disabilities. *Education and Training in Mental Retardation and Developmental Disabilities*, 36, 3–15.

Patterson, K. (2005, Winter). What classroom teachers need to know about IDEA '97. *Kappa Delta Pi Record*, pp. 62–67.

Rathbone, C. (2005). A learner's bill of rights. *Phi Delta Kappan*, 86, 471–473.

Torres, M. S., Jr., & Chen, Y. (2006) Assessing Columbine's impact on students' Fourth Amendment case outcomes: Implications for administrative discretion and decision making. *NASSP Bulletin*, 90, 185–206.

Trerise, R., & Johnson, C. (2002). Are you going to school in that? *Principal*, 81, 60–61.

SUICIDE

Human beings engage in a wide range of self-injurious thoughts and behaviors. These range from risk taking (e.g., skydiving), through suicide ideation, instrumental suicide-related behavior (e.g., suicide threats) to suicidal acts (attempted and completed). Suicide involves self-chosen behavior intended to bring about one's death in the shortest term. Each year, at least half a million people worldwide end their own lives. Difficulties in identifying unequivocal evidence of intent underpin the generally accepted view that official statistics underestimate the true suicide rate by an unknown amount. The presence of a note or letter often provides the clearest evidence of intent, but less than 30% of those who die by their own hand leave such a communication. Differing legal and ascertainment procedures account for some of the variation in national suicide rates. There may also be social pressures to underreport, especially in the absence of unequivocal evidence of intent, because of the stigma associated with suicide. In doubtful circumstances, attributing the cause of death to something other than suicide may lessen the pain for family and

friends. Defining attempted suicide is equally complex because the behavior is characterized by several dimensions, including the lethality of the method used, the level of medical injury inflicted, and the level of suicidal intent. In North America, evidence of intent is included in the definition of attempted suicide. In Europe, evidence of intent is of lesser importance in acknowledgment of the fact that divergent intentions may lead to a suicide attempt. Definitional variations contribute to an unknown extent to discrepancies among research findings on suicide and attempted suicide in different countries.

As a cause of death, suicide is one of the least understood. Those who attempt or complete suicide share a number of general characteristics or risk factors, including major depression, alcohol dependence, substance abuse, antisocial behavior, childhood abuse, and personality disorder. The majority of risk factors for suicidality (the occurrence of suicidal thoughts or behaviors) have been identified through retrospective studies, and the false-positive rate—the probability of wrongly inferring a link between a particular factor and suicide risk—is very high for this method of inquiry. This is because retrospective inquiries are prone to hindsight bias: Knowing the outcome of an event increases the post hoc estimate of the perceived likelihood of the outcome. Given an account of a person's life history and told that the person committed suicide, people are more likely to infer links between factors in the person's biography that might explain the cause of death than are those who are given the same account and asked to predict the likelihood of suicide. Knowing the outcome approximately doubles the likelihood of a perceived connection between an event in the person's life history and his or her cause of death.

Genetics and Family History

The traditional view of suicide is that it is the product of a profoundly disturbed mind: People who want to kill themselves are mentally ill and in need of treatment and care. Supporting evidence for this view can be traced to some of the earliest studies of suicide, conducted in the 1920s, which were based on interviews with the relatives of people who had killed themselves. A history of poor mental health was identified as a significant factor in the great majority of cases. However, an alternative perspective cautions that the presence of mental illness does not fully

explain suicide: The vast majority of people who suffer a psychiatric disorder neither think about nor attempt suicide. Later studies showed that a family history of suicide contributes about a twofold increase in suicide risk, even when controlling for family psychiatric history. The fact that a family history of mental illness and family history of suicide are separate risk factors prompted two lines of investigation: the first on the potential contribution of genetic factors and the second on the role of family dynamics.

The first studies of the role of genetics were conducted with samples of twins. Twins share the same family and social environment for the early part of their lives and have higher levels of closeness both in terms of the number of years spent together before leaving the parental home and in the frequency of contact afterward. Twins have a reduced risk of suicide, which supports the view that strong family ties reduce the risk for suicidal behavior. However, if a monozygotic (identical) twin commits suicide, the surviving sibling is at significantly greater risk of suicide compared to the level of risk for a dizygotic (fraternal) twin who loses a sibling to suicide.

Early studies indicated that the genetic factors related to suicide appeared to represent a genetic predisposition for the psychiatric disorders associated with suicide. Later studies sought to determine whether there may be an independent genetic component for suicide, and to clarify how genetics might be linked to other factors implicated in suicide, particularly major depression, childhood physical abuse, social phobia, alcohol dependence, and behavioral disorder. There is no evidence, and no suggestion that there may be evidence, of a specific gene for suicide, but several candidate genes for the transmission of suicide risk have been identified. These genes are linked with the neurotransmitter serotonin. The predominant effects of serotonin in the brain are inhibitory. Irregularities of the serotonin system have been implicated as a causal factor in violence and impulsive aggression, and it is well established that the serotonin system is involved in substance abuse disorders, especially alcoholism. Numerous abnormalities, most of which reduce the inhibitory effects of serotonin, have been found in the serotonergic system of those who attempt or complete suicide. Current estimates suggest that genetic factors account for between 30% and 50% of suicidality. It is highly likely that other, as yet unknown, genetic factors mediate the risk for suicidality independently of the genetic factors responsible for the heritability of the major psychiatric conditions associated with suicide.

Studies of familial factors have focused on the parent–child transmission of suicide risk. Some studies suggest a six-fold increase in risk for suicide attempt in children of a parent who has attempted suicide compared with families in which a parent has died through other causes. The transmission mechanism is not clear, but it seems likely that an attempted or completed suicide has the potential to profoundly influence the interpersonal dynamics of the family. Suicide-bereaved children generally come from families with a history of psychopathology and substantial family disruption. Families of suicidal young people are often described as disorganized, unstable, and rigid—characteristics that are also associated with the suicidal person. Although it is clear that suicide can negatively affect family dynamics, the possibility of a reciprocal interaction has not been studied in any detail. The occurrence of a suicide within a family increases the likelihood that other family members and friends will incorporate this into their repertoire of problem-solving behaviors. The peers of adolescents who have a friend who completed suicide have an elevated risk of suicidality relative to peers who have had no such experience.

Mental Illness

A psychiatric illness is a well-established risk factor for suicide across the life span. High suicide rates have also been linked with mood disorders, particularly bipolar disorder (sometimes referred to as manic-depressive disorder). About 30% of those with a history of bipolar disorder will commit suicide compared to about 12% of those suffering major depression. Personality disorders such as antisocial personality disorder or, in the case of younger people, conduct disorder are also linked with suicidality. Personality disorders involve a long-standing pattern of aggressive behavior, a reckless and impulsive disregard for the well-being of others, and contempt for social rules and norms. One explanation for the link between antisocial personality disorders and suicidality implicates a combination of impulsive aggression and anger, which are partly caused by deficits in the serotonergic system. Another explanation for the association between antisocial personality disorder and suicide points to the cumulative effects of repeated instances of self-harm, both

direct, as through immediate physical self-harm, and indirect, such as engaging in wantonly reckless actions. Repeated self-harm is strongly related to higher risk for eventual suicide.

Substance use disorders also confer risk for suicidality. People who are alcohol dependent with a history of suicide attempts tend to be more aggressive and are characterized by their reckless impulsivity. These characteristics act in combination to significantly increase the pathogenesis of suicide attempts in those suffering alcoholism. Similarly, the social profile of narcotics abusers is one of predominant unemployment, low educational levels, social isolation, divorce, repeated incarceration, high rates of parental alcoholism, and general psychopathology. A diagnosis of schizophrenia has long been associated with increased risk for suicide and suicide-related behaviors. Forty percent of those diagnosed with schizophrenia attempt suicide at some time in their lives. More than one third of those diagnosed will die by their own hand, a rate comparable to that for those diagnosed with mood disorders and more than 20 times higher than in the general population. The most predictive factors for suicidality among those suffering schizophrenia are the number of lifetime suicide attempts, the number of hospitalizations within the 3 years preceding the suicide attempt, a current or lifetime history of substance abuse, and major depression. The evidence for the significance of auditory delusions as a risk factor is not clear-cut. Some studies suggest that delusions may be significant either because patients hear voices directing them to suicide or because the interminable presence of auditory hallucinations becomes an unbearable stressor.

Social Isolation

Social isolation, a state in which interpersonal contacts and relationships are severely disrupted or non-existent, has been consistently related to suicidal behavior. The French sociologist Émile Durkheim was the first to propose and test the theory that suicide results in large part from social marginalization. He defined four types of suicide: egoistic, altruistic, anomic, and fatalistic. *Egoistic suicide* refers to circumstances in which a person has poor or nonexistent social support and feels little connection with the people living around him or her. *Altruistic suicide* denotes circumstances in which a person is overly integrated and comes to regard suicide as a social or religious duty, such as in hara-kiri. *Anomic suicide* is

associated with rapid social and economic change, which can cause an acute sense of alienation as might occur with the loss of a job and the absence of alternative employment opportunities. *Fatalistic suicide* occurs when people feel they have lost control of their lives and lose a sense of influence over what might happen to them in the future. Social isolation is detrimental to mental health, whereas the presence of a strong social network is a protective factor against mental illness and suicide.

Studies of adolescents show that those with a history of suicide attempts often chose not to seek support from others and did not talk with anyone about their thoughts during the period of suicidal ideation. The process of exploring one's sexual identity can be particularly stressful at a personal level and can fracture social relationships. Gay, lesbian, and bisexual adolescents have a rate of suicidal behavior—two to six times greater than that of their heterosexual peers. The inflated risk is due in part to the social exclusion that may accompany a minority sexual identity that has the potential to carry a social stigma.

Abuse

Physical or sexual abuse increases the risk for suicidal behavior. Adolescents with a history of sexual abuse are 12 times more likely to self-harm and nearly 50 times more likely to make a suicide attempt than peers without a sexual abuse history. Childhood sexual abuse is linked with suicidality through hopelessness and depression. Among those who have suffered sexual abuse, hopelessness is associated with high suicide risk only, whereas depression is associated with high suicide risk and actual attempts. Depression is more strongly associated with high suicide risk in girls than in boys.

Hopelessness

Hopelessness refers to a set of cognitive schemas whose common denominator is negative expectations about the future and impoverished positive beliefs in the likelihood of future success in any area of life. Hopelessness has a role in both causing and maintaining depression, and it is a significant predictor of suicide. Theories of suicidality that emphasize the importance of hopelessness contend that suicidal people have strong feelings of being a burden on others and of "not belonging," and they are specifically

hopeless about the prospects for these core areas of relationship management and development. These theories are based on evidence suggesting that people gradually acquire the ability to inflict lethal self-injury through prior experience of deliberate self-harm, which in turn is encouraged by a tendency to behave impulsively underlaid by serotonergic abnormalities. The potential to commit suicide will not be released unless the desire for death is motivated by a strong sense of perceived burdensomeness coupled with a sense of failed belongingness. These two factors, burdensomeness and belongingness, clarify why hopelessness, social isolation, and the psychiatric disorders for which they are associated features comprise a cluster of potent risk factors for suicidal behavior.

Burdensomeness and failed belongingness are linked to a depressogenic attributional style. Traditionally measured along three dimensions (internal–external, stable–unstable, global–specific), those who are suicidal tend to attribute negative life events to causes that are internal or personal to themselves, stable, and global. This pattern of thinking can be particularly invidious when applied to interpersonal experiences, such as the break-up of a long-term relationship. Adolescents bereaved through the suicide of a parent may believe they played a contributory role (an internal cause) either by their own actions or through their failure to notice signs of their parent's troubled mind. They may perceive their role in the suicide as an expression of some intractable personal trait (stable) and find confirmation for this in selective memories of other relationships in which they rejected, or were rejected by, others (a global feature of their interpersonal relationships). This pattern of thinking can be associated with an extended phase of negative affect and cognition, leading to decrements in problem solving and an increased sense of hopelessness, which in turn elevates their suicide risk. In general, there is still a taboo surrounding suicide, and those bereaved by suicide will have fewer opportunities to share their grief with others. Constricted problem solving may lead some to conclude that their own suicide would be a logically appropriate way of permanently removing the gap between themselves and their deceased parent.

Perfectionism

Perfectionism refers to a tendency to set unrealistically high standards and goals and to be excessively self-critical. Normal perfectionists set very high standards but understand that it is not always possible to achieve these, whereas neurotic perfectionists are more excessive. They are overly concerned with minor mistakes and tend to regard a task as a personal failure if it contains the merest hint of error. Driven more by a fear of failure than a desire for success, neurotic perfectionists constantly fall short of their own standards. This negatively affects their sense of self-efficacy, further exacerbating their sense of failure. They may also believe that they have failed to meet the expectations of others, which contributes to a growing sense of worthlessness. Neurotic perfectionism is linked with eating disorders, alcoholism, and suicidal ideation, and it increases the risk of suicide as an apparent solution to a life of unremitting failure.

Problem Solving

Impaired problem solving is a well-established suicide risk factor. Those at high risk tend to endure impairments in their abilities to solve a range of social and interpersonal problems. This is often most noticeable in difficulties experienced when identifying and formulating appropriate solutions to familiar social problems. Suicidal people generate fewer solutions, and often, their solutions are less relevant to the problems they are trying to solve. The suicidal also tend to generate overly general autobiographical memories—summaries of experiences rather than details—and take longer to recall positive memories. This is explained in terms of a *mnemonic interlock;* the suicidal person is fixed at an intermediate level of memory recall such that he or she can access general but not specific memories. Solving personal and interpersonal problems entails recalling similar dilemmas and routes to satisfactory solutions. Difficulties accessing specific events from long-term memory are likely to diminish one's problem-solving capacity. In effect, suicidal individuals engage in efforts to solve dilemmas burdened by an information-processing handicap.

Diathesis-Stress Theories

Diathesis refers to propensity or predisposition, and for suicide, it relates to the risk factors for suicidality. The risk of suicide is not directly related to objective indicators of the severity of a particular stressor, such as the gravity of mental disorder, but instead depends on variations in the diathesis. Diathesis-stress theories

explain suicidality in terms of an inherited susceptibility to suicide combined with a particularly stressful life event and a diminished set of resources, such as weak social networks and poor problem-solving skills. The varieties of diathesis-stress theory have in common an emphasis on the contributions of depression, hopelessness, attributional style, and problem solving, and they differ principally in the relative importance given to each of these factors.

Entrapment and Escape Theory

Escape theory contends that the principal motive of suicide is to flee the painful self-awareness of particular interpretations or implications about one's self. Escape theory regards suicide as the final step in a series of causally related events that begins with a particularly stressful experience, or series of experiences, that is appraised as a blameworthy personal failure. This leads to deeply negative affect from which the person wants to escape coupled with a sense of cognitive and emotional numbness. The resultant numbness culminates in disinhibition toward a wide range of behaviors, including self-injury, which increases suicidal risk. Escape theory has been elaborated through a linkage with research on "arrested flight" or entrapment in animals. According to this view, suicidal behavior is a response to a stressful situation that has three components: (a) the stressful situation causes feelings of defeat or rejection; (b) the person wishes to escape from the situation and concludes that there is no escape; and (c) he or she cannot identify rescue factors, such as social support, to alleviate the crisis. In combination, these factors are thought to activate a psychobiological "helplessness script" that facilitates an impulse to escape through suicide. Whether or not a suicidal solution is preferred is determined by other factors, including modeling effects—whether the person knows someone who has used suicide as a solution—and the availability of lethal methods. A person's judgments as to how stressful and escapable the situation is, and how much support is available, are affected by memory and problem-solving skills.

Career Theory

The notion of "suicidal career" is based on the concept that suicide derives from an inability to accept or cope with what it means to be human. The "human condition" encapsulates the idea that all humans share the same fundamentals, including a limited life span, restrictions on the ability to engage in meaningful work, love relationships, health, and strength. The harshness of the human condition varies from person to person, and most people can deal with the harshness of life without resorting to suicide. However, some of the methods used to cope with being alive, such as substance abuse, are essentially self-destructive, and a person's repertoire of problem-solving skills can become progressively constricted. For some, suicide comes to be regarded as rational because, with diminishing problem-solving skills, it seems to offer a way to solve the problems that are an unavoidable part of being human. Evidence of a suicidal career can often be found in the histories of people who have completed suicide: Life has been very difficult, anxiety provoking, and painful. Many will have used nonsuicidal means of dealing with their problems, including some self-injurious methods, that have proved unsuccessful, and suicide appears to offer a solution. With hindsight, it is usually possible to identify sequences of events and problem-solving methods that suggest a career toward suicide.

Suicide Prevention

Interventions designed to prevent suicide focus on enhancing the protective factors for suicide and reversing or reducing known risk factors. The strongest protective factors are effective clinical care for mental, physical, and substance use disorders; easy access to a variety of clinical interventions and support for help-seeking; restricted access to highly lethal means of suicide (e.g., guns); strong connections with family and friends; support through ongoing medical and mental health care relationships; skills in problem solving and conflict resolution; and cultural and religious beliefs that discourage suicide and support personal well-being and development.

The goal of suicide prevention is to break the causal chain that leads to self-injury. Early identification of those at risk is an important part of any preventive strategy. Primary health care personnel may not be familiar with recent advances in the identification of people with risk factors for suicide, and the provision of professional development education programs aimed at the early identification of those at risk can contribute to a reduction in rates of attempted and completed suicide. New drugs for the treatment of bipolar and schizophrenic disorders have a more specific therapeutic profile than those used previously

and appear to improve patients' adherence to treatment and reduce the likelihood of suicidal behavior.

In suicide prevention, *targeting* refers to a focus on modifying something that is causally related to suicidality within a specific population. Prevention targeting is structured at different levels and in different stages. A commonly used framework for describing intervention levels distinguishes between *indicative*, *selective*, and *universal. Indicative* or *indicated interventions* are highly targeted and usually involve identification, treatment, and skill building among individuals at high risk. *Selective interventions* are targeted at groups whose members are likely to include some high-risk individuals. The focus is on screening and group prevention activities. Peer support programs for students with a number of risk indicators would be an example. *Universal interventions* are targeted at communities and may include media or educational campaigns and other broad-population-based prevention strategies.

For each of the levels, there are three stages of suicide prevention: primary, secondary, and tertiary. These correspond to before suicidal behavior occurs, as suicidal behavior occurs, and after suicidal behavior occurs, respectively. Primary intervention targets the risk factors for suicidality such as alcohol and drug abuse or programs that prevent social isolation, such as bullying prevention. Secondary intervention focuses on the early detection of suicidal ideation, or planning, and appropriate referral. Tertiary prevention targets intervention following self-injury, such as evaluating the effectiveness of a therapeutic treatment and the provision of appropriate support services. A wide range of suicide prevention efforts are ongoing around the world, although there is very little evidence as to the efficacy and effectiveness of any of them. Suicide prevention strategies continue to be developed from the principle that doing something is better than doing nothing. However, hazards may result directly from prevention efforts. For example, vulnerable adolescents may be distressed through exposure to certain suicide prevention education curricula. A growing appreciation of the potential for negative impacts has added to a sense of urgency to understand which interventions are most effective.

Curriculum programs, staff in-service training, and screening programs are several school-based interventions that have been implemented. The aims of curriculum programs are to raise student awareness of the risk factors for suicide, teach students to recognize possible signs of suicidal behavior in order to assist others, and provide students with information about various support resources in the school and the wider community. The efficacy of these programs has been criticized because they produce gains in knowledge about suicide, but there is no evidence that they have an impact on suicidality. Staff in-service training programs include many of the features of curriculum programs. Typically, they also provide training in the identification of potentially suicidal students, discussions of sample cases, and information regarding referral procedures. These programs raise knowledge levels and referral practices, but their efficacy in reducing suicidality has yet to be determined. Screening programs are schoolwide assessment regimes designed to identify potentially suicidal students using a two-stage procedure. First, all students complete a self-report screening measure in order to identify potentially suicidal students. Second, those students who score above clinically significant levels on the initial screening are offered interviews with an appropriately qualified school-based professional with a view to obtaining a more precise assessment of suicide risk. Direct assessment of this kind may be more effective and efficient than curriculum or staff training programs. Early evaluations are limited in number, but the outcomes are promising.

Noel Sheehy

See also Aggression; Child Abuse; Eating Disorders; Failure, Effects of; Self-Efficacy

Further Readings

Borowsky, I. W., Ireland, M., & Resnick, M. D. (2001). Adolescent suicide attempts: Risks and protectors. *Pediatrics*, *107*(3), 485–493.

Gould, M. S., Greenberg, T., Drew, V., & Shaffer, D. (2003). Youth suicide risk and preventive interventions: A review of the past 10 years. *Journal of the American Academy of Child and Adolescent Psychiatry*, *42*(4), 386–405.

Hawton, K., & Van Heering, K. (Eds.). (2002). *The international handbook of suicide and attempted suicide*. Chichester, UK: Wiley.

McLaughlin, C. (2007). *Suicide-related behaviour: understanding, caring and therapeutic responses*. Chichester, UK: Whurr.

O'Connor, R., & Sheehy, N. (2000). *Understanding suicidal behaviour*. Leicester, UK: BPS Books.

Spirito, A., & Esposito-Smythers, C. (2006). Attempted and completed suicide in adolescence. *Annual Review of Clinical Psychology*, *2*, 237–266.

Williams, J. M. G. (2001). *Suicide and attempted suicide*. London: Penguin.

T

Tell me and I'll forget; show me and I may remember; involve me and I'll understand.

—Chinese proverb

T SCORES

T scores are area-normalized scores standardized to have a mean of 50 and a standard deviation of 10. They are like normally distributed *z* scores except that they have been subjected to a linear transformation that changes the mean of the distribution from zero to 50 and changes the standard deviation from 1.0 to 10. As a result, *T* scores are always positive, usually ranging from about 20 to 80, although more extreme values are possible. Ordinarily, *T* scores are reported without a decimal point; that is, they are rounded to whole numbers (48.4 becomes 48, 48.6 becomes 49).

T scores were first described by McCall and named in honor of three famous psychologists of the era whose last names began with T. Since that time, *T* scores have provided the metric for a number of psychological and educational tests. Although McCall defined *T* scores as being normalized, thereby giving them many of the advantages of normal-curve *z* scores without some of the disadvantages, subsequent users have often failed to incorporate the normalization step and have simply transformed ordinary *z* scores into a scale with a mean of 50 and a standard deviation (*SD*) of 10 and called them *T* scores. Thus, when *T* scores are mentioned as a metric for a test, it is important to determine whether the normalization step has been taken.

Properly constructed *T* scores cannot be computed directly from the raw scores unless the raw scores are themselves normally distributed. This is because any linear transformation, such as computing *z* scores, does not change the shape of the original distribution. If the raw score distribution is positively skewed, the *z* scores and any transformations of those *z* scores will also be positively skewed to the same degree.

The determination of area-normalized scores requires the use of a table of the normal distribution such as is found in almost any statistics book and some measurement texts, and it always involves as a first step the determination of normalized *z* scores. The process is as follows:

Make a frequency distribution and cumulative frequency distribution of the raw scores.

Determine the percentile rank of each raw score. The formula is

$$PR_j = \frac{cf_{j-1} + .5f_j}{N},$$

where

PR_j is the percentile rank of raw score *j*,

cf_{j-1} is the cumulative frequency for the next lower raw score,

$.5f_j$ is half of the frequency for raw score *j*,

N is the total number of cases.

Once the percentile ranks have been determined, look up the z score that has the corresponding percentile rank.

a. For percentile ranks less than 50, this will be the negative z score, which has an area beyond itself equal to the percentile rank. For example, if the percentile rank of a raw score is 16, the z score that has 16% of the distribution below it is -1.00.

b. For percentile ranks greater than 50, subtract 50 from the PR and look up the z score that has the remaining percentage between itself and the mean. For example, if the percentile rank of a raw score is 84, we would look up the z score that has 34% of the distribution between itself and the mean. This z score is $+1.00$. (Occasionally, you may find a table of the cumulative normal distribution in which the relationship between percentile ranks and z scores can be read directly.)

c. You now have a set of z scores that has been forced into a normal distribution by being assigned values corresponding to the percentile ranks of the normal distribution. To obtain T scores, multiply each z score by 10, then add 50. That is,

$$T_j = 10(Z_j) + 50.$$

Area-transformed z scores can be converted into any arbitrary metric (e.g., mean of 100, SD of 20; mean of 500, SD of 100) using the following formula

$$A_j = SD_A Z_j + M_A,$$

where

A_j is the transformed score of raw score j in an arbitrary scale,

SD_A is the standard deviation of the chosen arbitrary scale,

Z_j is the area-transformed z score for the particular raw score,

M_A is the mean of the chosen arbitrary scale.

It is possible to apply any linear transformation to normalized z scores. For example, essentially all intelligence tests report their scores as normalized scores with a mean of 100 and a standard deviation of 15. Other commonly used scales apply a mean of 100 and an SD of 20 (Armed Forces tests) or a mean of 500 and an SD of 100 (SAT scores). When a test publisher has applied an area normalization to the standardization data for a test, a table of equivalent scores will be provided to convert raw scores directly into normalized standard scores in the publisher's chosen metric.

These transformed scores are identical to T scores except for the final metric. Normalized scores have the advantage over raw scores or simple linear transformations in that there is a link between the normalized score and percentiles, making normative interpretation easier. T scores and other transformations of normalized z scores have the advantage over percentile ranks that they can be treated more plausibly as representing an interval scale of measurement.

It is appropriate to perform an area-normalizing transformation whenever one has reason to believe that the underlying property as it exists in nature has a normal distribution and that one's measurement process is responsible for the non-normality. For example, there is reason to believe that intelligence is normally distributed in the general population, but if the test being used contains items that cluster at an ability level at one extreme of the distribution, a skewed distribution would result. In such a case, area normalization would be appropriate. However, some characteristics, such as reaction time, are likely to have skewed distributions in nature. When this is the case, area normalization is not appropriate, but a nonlinear transformation of the original scores such as a logistic or arcsin transformation would be appropriate and may yield normally distributed scores as a precursor to further statistical analysis.

Robert M. Thorndike

See also Descriptive Statistics; Inferential Statistics; Statistical Significance

Further Readings

McCall, W. A. (1922). *How to measure in education.* New York: Macmillan.
Thorndike, R. M., & Thorndike-Christ, T. (2008). *Measurement and evaluation in psychology and education* (8th ed.). Upper Saddle River, NJ: Prentice Hall.

TEACHING STRATEGIES

Teaching strategies are the procedures, processes, activities, and tools used to assist in learning. These strategies encompass a wide range of approaches and

actions and are situated across a variety of contexts. Instructors and teachers use teaching strategies to enhance learning from preschool through college settings. Many other professionals and laypersons use teaching strategies, from parents teaching a child how to ride a bike, to physicians teaching interns about medicine, to mechanics teaching apprentices how to repair a car.

There is a wide range of teaching strategies available to those involved in helping others learn. The type of teaching strategy used will vary according to the students' characteristics; the learning task; the situation; the context demands; and the knowledge and skills of the "teacher," whether that person is a parent, professor, or mechanic. Most effective teachers and instructors will use a variety of teaching strategies, after carefully considering all of the factors that affect instruction.

In the following sections, a brief history of the development of teaching strategies is provided, as well as a description of a variety of teaching strategies, including instructor- and student-centered options. New developments in technology and teaching are also discussed.

Historical Contributions to Teaching Strategies

Interest in effective teaching strategies has been a focus of discussion and research for hundreds of years. Perhaps one of the earliest teaching strategies was described by Plato in *Socratic Dialogues.* Plato described what has come to be known as the *Socratic method* of teaching, which is a variant of a strategy used by Socrates. Still used today as a teaching strategy, the Socratic method is a dialectic method of teaching that involves dialogue and questioning, emphasizing the exchange of ideas and suppositions that then transforms knowledge itself.

Many significant contributions to our knowledge of teaching strategies have been made in the past two centuries, continuing today with exciting new developments (e.g., brain research). John Dewey was an American psychologist, philosopher, and educator whose contributions to the field of education continue to be evident today in many teaching approaches. Although he wrote extensively on many different topics (e.g., democracy, human conduct, logic), his contributions to the field of education focused on teaching approaches. Dewey felt that content area

subjects should be integrated when taught and that experiential education—learning by doing—was the most effective way to teach. He also believed that teachers should emphasize critical thinking as opposed to memorizing facts, and that problem solving and inquiry were two concepts that should be embedded in instruction.

Ivan Pavlov's work in the 1920s and the research of Edward Thorndike and B. F. Skinner have also had a profound effect on the development of teaching strategies. Pavlov studied stimuli and responses in an effort to understand learning and is famous for discovering *classical conditioning* involving involuntary responses. Thorndike and Skinner expanded Pavlov's research by examining learned (voluntary) behaviors and were major contributors to the research on operant conditioning. They examined the antecedents and consequences of behavior. Thorndike discovered that any behavior is more likely to be repeated if it results in a positive outcome (the law of effect). Skinner studied the antecedents of behavior, the behavior itself, and the consequences of behavior and is generally viewed to be the founder of behavior modification. These researchers contributed to the foundation of teaching strategies based on the principles of behaviorism.

Jean Piaget and Lev Vygotsky contributed to the development of teaching strategies based on cognitive and social interaction learning theories. After studying how children learn, Piaget theorized that humans go through four stages of cognitive development: sensorimotor, preoperational, concrete operational, and formal operational. How one thinks becomes more sophisticated at each stage, and therefore, Piaget posited that teaching strategies would need to be adapted to each stage. Although Piaget found that social transmission (learning from others) was an important aspect of learning, he did not focus on language and communication as the primary vehicle for learning.

As Vygotsky examined learning, he theorized that not only were social interactions a key to learning, they created learning. His research on language, culture, and cognitive development was a significant contribution to the sociocultural theory of learning. Vygotsky felt that through interactions with those around them, children co-construct knowledge. One key element in the group co-construction of knowledge is having less capable students work with more capable students or adults. The more capable person (e.g., teacher) can enhance the less capable students' learning by providing scaffolding; that is, support during the learning task.

Scaffolding is a common term in teaching and learning and refers to visual and verbal prompts and supports that facilitate learning. Many constructivist teaching strategies and approaches to teaching are based on Piaget's and Vygotsky's ideas and theories.

Building on the social-cognitive theories of learning, Albert Bandura, a Ukrainian-Canadian psychologist, developed the social learning theory. His research revealed the power of modeling as a tool for learning. Bandura found that learning occurs from observing others and from being reinforced for learning successes. His research also focused on self-efficacy, or the belief that one has about one's ability to perform satisfactorily. These beliefs affect the learner's perception of a task, as well as the motivation to attempt a task.

Jerome Bruner, an American psychologist, focused on how students understood the conceptual underpinnings of knowledge. He was interested in how to structure instruction so that conceptual understanding was achieved. Whereas Piaget felt that learning was determined by what stage one occupied, Bruner's research concluded that learning was not dependent on one's stage (or age), but on environmental and experiential factors. Therefore, the way in which instruction was organized and what was emphasized during instruction were keys to improving learning. Inquiry and discovery learning teaching strategies grew out of Bruner's work on using inductive reasoning to improve students' deep understanding of concepts and structures of knowledge.

Research has also focused on how concepts and ideas are taught, as opposed to how the learner discovers or constructs knowledge. Whereas Bruner focused on inductive reasoning as a path to learning, David Ausubel, an American psychologist, focused on deductive reasoning to enhance learning. Ausubel's research resulted in the development of more direct teaching strategies. Expository teaching strategies that emphasize using verbal information to help learners understand concepts, particularly the use of advance organizers, have grown from Ausubel's studies. Advance organizers are tools used by instructors to help learners organize new and existing information. Constructed and presented before students are expected to learn new concepts, they provide an external scaffold upon which students can develop internal schemas to classify and construct knowledge.

Another educational researcher who focused on the organization and delivery of instruction, Siegfried Engelmann, investigated methods of teaching that would enhance the learning of students challenged by poverty and educational barriers. Engelmann developed approaches to instruction not based on Piaget's idea of stages or on the discovery learning ideas of Bruner but rather on how to set up learning experiences so that all students are actively involved, they are given many opportunities to respond, and during which teachers continually assess student performance and provide feedback. After conducting several large-scale research projects, Engelmann developed the direct instruction method of teaching based on effective instructional components including homogeneous small group instruction, carefully designed instruction, active responding, and frequent feedback and correction.

From this very brief overview of the historical foundations of teaching strategies, several themes emerge. There are differences in the ways learning is perceived, from those that focus on stages of development to those that focus on the interactions between the learner and the type of instruction. These differences are reflected in the diversity of teaching strategies used today.

Examples of Effective Teaching Strategies

Most teaching strategies have emerged from the research done on learning. The logic is that if learning occurs under certain circumstances, such as in a group dialogue or from observing, then teaching strategies should match those circumstances. As a result, teaching strategies have been classified according to a wide range of criteria. Student-centered, inquiry-oriented, constructivist, direct instruction, and cognitive are some of the categories used to describe teaching strategies, all of which emanate from philosophical and/or research-based perspectives on learning and teaching. This discussion includes two broad categories of teaching strategies: instructor-centered and student-centered. In instructor-centered teaching strategies, the instructor takes an active and engaged role throughout the teaching process. Of course, all instructors are most effective when they are actively engaged regardless of the teaching strategy being used. However, a key distinction between instructor- and student-centered strategies is the direct and continual involvement of the instructor *throughout* the learning process in instructor-centered strategies. In the category of student-centered teaching strategies, the instructor may take on the role of

observer, guide, facilitator, or even mediator during the learning process, leaving aspects of the learning to occur within and among the students. Student-centered teaching strategies involve similar planning and preparation efforts on the part of the instructor, but less direction and control during the learning process.

Instructor-Centered Teaching Strategies

Expository Teaching Strategies

Expository teaching strategies involve the instructor verbally leading and engaging students in the learning process. These strategies include the traditional lecture method or smaller, mini-lectures/presentations. In the traditional lecture teaching strategy, certain elements are associated with positive learning outcomes: gaining attention and generating anticipation, presenting the objective of the lesson/lecture, stimulating background knowledge and review of previously addressed material, verbal explanation of the concepts or ideas, providing examples that lead to learning, providing practice opportunities and feedback, assessing the learning, and summarizing.

When using expository teaching strategies, the instructor must first gain the learners' attention and generate interest in the topic. Many techniques can be used to achieve this, such as posing a problem of interest or asking a provocative question—both of which should be of interest to the audience and related to the topic. Concrete or visual items can be presented, such as a model, an example, or a picture to stimulate interest. Next, the instructor provides the objective for the lesson so that students can understand the ultimate purpose of the lesson and also follow a mental route to that end. The instructor then connects that objective or topic to previously learned material and/or to the students' background knowledge.

During the verbal explanation of the topic, the instructor can, of course, talk. But most instructors recognize the power of using additional strategies at this point. For example, visual aids such as graphic organizers, semantic maps, charts, curriculum webs, graphs, video clips, pictures, and diagrams all enhance the learning process. A clear sequence that flows from concept to concept, reconnecting previously discussed ideas and continually offering examples and clarifications, enhances learning. The instructor then offers practice opportunities for students, such as writing exercises, lab experiments, math problems, questions, or case study examples. Practice opportunities can range from 1-minute writes, to short questions posed for limited discussions among peers, to more complex and involved practice items that may or may not be accomplished in one lesson. Students can be assigned to or form self-chosen groups to accomplish the practice. During practice, the instructor provides feedback and error correction. If a concept is very new to students, then the instructor offers more frequent feedback and correction. During practice, the instructor should be assessing students' understanding. Assessment can also be accomplished through a question-and-response period led by the instructor or by evaluating the outcomes of the practice activities. At the conclusion of the expository lesson, the instructor summarizes by restating the lesson objective and reviewing the concepts addressed.

Mini-lessons/presentations are structured in the same manner as standard exposition. The difference is that the verbal explanation (i.e., the actual lecturing) comprises only a small proportion of the lesson and the student practice time is extended. Mini-lessons work well with younger students and those with attention difficulties, or when a longer lecture is just not possible due to time constraints.

Interactive-Expository Teaching Strategies

Similar to expository teaching strategies, the interactive-expository strategies are sequenced in the same manner, but additional elements such as interactive questioning (e.g., using the Socratic method), modeling, and high levels of student responding are embedded in the lesson. The Socratic method as it is used today can best be described as a discussion in which the leader (teacher) poses questions to students, and then responds to student input with additional questions designed to facilitate student engagement and active knowledge construction. Like Socratic methods, most interactive-expository strategies contain elements that resemble more of an interactive dialogue with the students than a traditional lecture. The instructor uses student responses and input during the lesson to adjust the pace and content of the exposition. Students are encouraged to participate in the dialogue by discussing topics with their peers, writing responses (e.g., the 1-minute write), posing questions, completing graphic organizers, constructing semantic maps, or paraphrasing content presented. In "instructional conversations," the instructor builds understanding among the students by mediating student dialogue so that the students are actually co-constructing their knowledge and understanding

through the conversations. The instructor poses questions, elaborates on students' comments, and redirects students' input to achieve the learning outcomes.

Modeling

Modeling is a powerful teaching strategy that can take many forms and be embedded into any learning experience. However, modeling a task (e.g., measuring volume), a procedure (e.g., making a chemical solution), or the production of an ideal product (a research paper) should contain verbal explanations that clarify and enhance learning. The verbal explanations do not require a lecture per se, but can be short explanations of what the instructor is doing and why, while emphasizing the expected outcome. Some refer to these short explanations as "think alouds." For example, a science teacher might model an entire scientific procedure from start to finish as he or she verbally explains his or her thinking, purpose, and expected outcome. An English teacher may model the process of writing a paper, starting with the outline, while explaining his or her thinking to the students. After modeling the process, the instructor may show an example of the product— a well-written paper—at which time the instructor points out the critical attributes of a quality product.

Modeling may be used to show students parts of a task, with student practice following each modeling session. If a task is complex and in-depth, and the students have little background knowledge, often the instructor will model components of the task, have students attempt that component, and then, based on how the students performed, either move on to the next modeling session or repeat the first modeling session.

Direct Instruction

Direct instruction teaching strategies include homogeneous small group instruction, high rates of student responses, frequent feedback from the instructor and error correction, brisk pacing, sequenced lessons, and instructor scaffolding. Small groups that consist of students working at the same level in a given area, such as word decoding, algebraic understanding, reading comprehension, or grammatical skills, are identified through assessment and formed for instruction. It should be noted that these are not long-standing groups, such as the traditional ability groups. These types of small, similar-skill-level groups will change as students become more proficient. Students may master concepts and move to different instructional groups at different rates.

In direct instruction, the lessons are sequenced carefully so that new skills and concepts are introduced only after students have mastered the prerequisite knowledge or skills. Students continue to practice mastered skills but are confronted with only a limited number of new concepts in each lesson. They are provided with carefully constructed examples and nonexamples of the concept or skill. While working with the small groups, the instructor follows a routine that includes research-proven effective teaching behaviors. The instructor continually asks for student responses during the lesson and assesses student understanding and mastery through those responses, providing feedback, practice, and error correction after each response. The pace of the lessons is brisk, but not hurried. Although the intensity of the lessons is high, with many student–teacher interactions, the lessons are not long. The idea is to intensively teach a skill or concept during the lessons and follow with student practice on those specific skills. Student practice occurs immediately following the lesson, with the teacher providing continuing feedback on students' performance. The students are also given extended practice, such as homework focused on skills that they have mastered.

Student-Centered Teaching Strategies

Although instructors will not be directly involved in all aspects of learning when using student-centered teaching strategies, they must plan extensively for these strategies to be successful. Instructor involvement may range from being a close guide throughout an activity, to an observer noting what students are discussing, how they are progressing, and if the learning objective is being accomplished. Many student-centered teaching strategies have developed from a constructivist view of learning that emphasizes co-construction of knowledge through students' verbal and experiential interactions.

Peer Learning Groups

Many instructors use student groups as a teaching strategy. These groups can take a variety of different forms, from homogeneous to heterogeneous, and be used for a wide range of learning objectives. Homogeneous groups can be skills-based with students at similar skill levels (e.g., all understand two-digit addition, but not two-digit subtraction) that are grouped together for instruction. Homogeneous groups can also be interest based (e.g., all students with an interest in nonfiction

books). Heterogeneous groups comprise students with dissimilar characteristics or interests (e.g., students ranging from first-grade to sixth-grade reading abilities grouped together or students with different interests in science in one group).

Skills groupings are homogeneous groups formed by assessing students' skill levels in specific content (i.e., reading) or depth of understanding of a particular subject (i.e., chemical reactions). Students are identified and assigned to groups based on those similarities. The tasks performed within these groups are designed to match the students' specific capabilities. Skills grouping is a teaching strategy that can be used with struggling as well as with highly capable students.

Interest groupings can be formed based on students' similar interests so that students can pursue a deeper understanding in an area of interest (e.g., marine animals, cars, astronauts). Although students may be at different skill levels in these groups, their homogeneous interests would allow them to work together on projects or assignments.

Perhaps the most popular heterogeneous grouping technique used as a teaching strategy today is the cooperative learning group. These groups are heterogeneous because the goal in the formation of the groups is to have a wide diversity of knowledge and skill levels represented. Setting up successful cooperative learning groups and tasks involves very specific techniques. The instructor must set up the learning task so that there is positive interdependence; that is, to achieve the learning outcome, all members of the group are dependent on each performing his or her assigned task or role. If an individual's task is not completed or a role is not fulfilled, that will affect the entire group's accomplishment. The roles and tasks can be assigned according to each member's abilities and skills. For example, if one member is unskilled at writing, that member may be the person who collects the needed information, but is not required to write.

Another aspect of effective cooperative learning is individual accountability, in which each member is not only accountable for a job or task, but also is accountable to the entire group. Group feedback forms on how each member contributed to the group project is one way to promote accountability. Cooperative learning group tasks are structured so that the process includes face-to-face interaction. If the task is set up so that students divide the jobs, individually complete the jobs, and then reconstitute the group, then the task is not designed for cooperative learning.

Student reflection and goal setting should also be an integral part of a cooperative learning task.

There are many different types of cooperative learning techniques, from "numbered heads together" to "jigsaw" techniques. Research has shown that cooperative learning groups improve understanding and achievement in all content areas; improve social interactions and social skills; and promote inclusion of a diverse set of abilities, interests, and perspectives.

Peer tutoring groups, another heterogeneous grouping strategy, involves matching more capable students with those who are less capable in a given area or skill. Peer tutoring groups can range from one-on-one tutorials to small groups in which a more capable student may teach three or four students who demonstrate lower skill levels or who are in need of more practice. Peer tutoring is strongly supported by research that shows it improves both the less capable and highly capable students' performance. The instructor must train students to effectively implement their assigned tutorial or teaching roles, provide feedback on the tutors' performance, and assist the peer tutors in solving ongoing instructional challenges.

Inquiry, Discovery, and Problem-Based Strategies

Teaching strategies for inquiry, discovery, and problem-based learning share similar components. They are all based on the idea that the instructor has facilitated student curiosity and interest in an area of study. The curiosity may have been piqued by a field trip to a museum, a book, or a lecture. The strategy involves providing lessons to build students' background knowledge and then providing them with a prompt, such as an interesting question (e.g., How were the pyramids constructed? How could we make cars better?) or an authentic problem (e.g., the aquifer is being reduced at a rate faster than it is being replenished) that encourages exploration, study, and research. Students usually work in groups to investigate their question, collect data, and develop conclusions. The instructor enhances the group process by providing guidance, asking additional questions, directing students to information and resources, and providing social skills support. With less capable students, students with little background knowledge, or those who have not mastered prerequisite knowledge, the instructor may have to provide more guidance, instruction, and assistance at every stage of the project.

New Developments

Although there are many new developments emerging from research on teaching strategies, such as brain studies and culturally responsive teaching, technology has recently had the greatest impact on how instructors teach and students learn. Better and more effective computer software programs designed to facilitate critical thinking, problem solving, and decision making, as well as provide practice on basic skills, are being developed. Projectors in the classroom, the Internet, personal response systems, online learning, blogging, Web-based learning, and electronic portfolios are all affecting instruction and new teaching strategies are being designed using technology applications.

Darcy Miller

See also Brain-Relevant Education;
Cooperative Learning; Direct Instruction;
Discovery Learning; Education; Expert Teachers;
Precision Teaching

Further Readings

Bruner, J. S. (1996). *The culture of education.* Cambridge, MA: Harvard University Press.

Dell'Olio, J. M., & Donk, T. (2007). *Models of teaching.* Thousand Oaks, CA: Sage.

Dewey, J. (1913). *Interest and effort in education.* Cambridge, MA: Houghton-Mifflin.

Joyce, B. R., Weil, M., & Calhoun, E. (2004). *Models of teaching* (7th ed.). Boston: Allyn & Bacon.

Mayer, R. E. (2008). *Learning and instruction* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.

Moore, K. D. (2005). *Effective instructional strategies: From theory to practice.* Thousand Oaks, CA: Sage.

Morrison, G. R., Ross, S. M., & Kemp, J. E. (2004). *Designing effective instruction* (4th ed.). Hoboken, NJ: Wiley.

Piaget, J. (1970). *The science of education and the psychology of the child.* New York: Orion.

Vygotsky, L. S. (1993). *The collected works of L. S. Vygotsky: Vol. 2* (J. Knox & C. Stevens, Trans.). New York: Plenum.

Woolfolk, A. (2004). *Educational psychology* (9th ed.). Boston: Allyn & Bacon.

Web Sites

Center for Applied Special Technology (CAST): http://www.cast.org

National Institute for Direct Instruction: http://www.nifdi.org

TEST ANXIETY

Test anxiety, a worried, restless, agitated distress that results from tests of performance or academic ability, affects everyone: from athletes, to students, to executives. Whether it is triggered by an upcoming speech, an impending athletic competition, a test of curriculum mastery, a college entrance exam, or a critical business decision, test anxiety can either enhance performance or hinder it, depending on whether the triggering event requires physical or intellectual performance. Unlike trait anxiety, which causes worry and distress over an extended period of time, test anxiety is a kind of state anxiety and happens only when one is in a specific situation requiring performance or evaluation. Test anxiety has become of particular concern to educators, students, and parents in the United States since the enactment of No Child Left Behind (NCLB). Internationally, there is heightened concern in recent years over academic testing of toddlers and teens cheating to gain admission into desirable colleges. In the United States, under NCLB regulations, schools face major upheavals if students are not passing standardized tests at the required levels. In what follows, the concept of test anxiety is explained, followed by an overview of test anxiety and the reason for concern within the framework of educational psychology and current standardized testing practices.

Understanding Test Anxiety

Sarason called the mental short-circuiting that results from test anxiety *cognitive interference.* When cognitive interference happens, emotional fear and uneasiness redirect logical, purposeful thought to distractions or sometimes reactions. Most people can recall experiencing the fight-or-flight response to a challenging or threatening situation. When cognitive interference occurs, thinking is replaced with avoidance or illogical choices that an individual would probably not make if he or she were able to think clearly. However, when test anxiety occurs, forethought disappears until the anxiety or fear diminishes. Then, when the anxiety reaction subsides, hindsight follows and the typical individual experiences that "Why did I do that?" remorse.

The emotional overriding of logical thought and memory retrieval has been visually illustrated through

medical imaging technology. Imaging pictures can show how the basal ganglia (the anxiety regulator) becomes overactive when an individual feels a test anxiety threat. The overactivity of the basal ganglia immobilizes thought processes. When an individual perceives a threat, the hypothalamus (the limbic system) provides an automatic, uncontrolled reaction to a perceived emotional or physical threat, acting as a circuit breaker between the prefrontal cortex (the center of purposeful thought in the brain) and the limbic (emotional center of the brain) systems. When the hypothalamus sends a signal, the limbic system goes into action. Emotions take over, and thinking and problem solving stop. Reactivation of the cerebral cortex (the thought processor) is not possible until the perceived threat no longer exists. The scientific evidence of imaging provided by those in the medical profession thus gives an objective explanation of test anxiety that can be easily understood when related in lay terms. Understanding that test anxiety happens and what causes it to happen can be the beginning of coping with test anxiety and overcoming it.

Test Anxiety in an Educational Framework

Students experiencing test anxiety can have different characteristics. Supon pointed out three types or categories of test-anxious students. Students in the first category lack the study skills to adequately prepare for tests and therefore lack the knowledge to perform well on tests. Students in the second category have the study skills necessary to prepare for the test, but have a fear of failure that makes them unable to perform successfully in test situations. The result, according to Meichenbaum, is a general lack of self-confidence. Students in the third category of test-anxious students believe that they have the study skills, but they do not. As a result, students in this category do not adequately prepare for the test, and the poor preparation for the test causes anxiety. Additionally, one is not simply test anxious or not test anxious. According to McDonald, there is a continuum of impairment, rather than test anxiety being either present or not present. The concern is that not all students may have the skills necessary to achieve passing scores on standardized tests when it is required that students at the same grade level take the test at the same time. A failure to allow enough time for individual students to master the

standard curriculum could be predicted to result in an increase in test anxiety for both teachers and students.

Empirical studies have shown that the anxiety experienced by teachers has a significant impact on the test anxiety that is experienced by students. If teachers experience stress and anxiety, students can be expected to experience stress and anxiety. Conversely, if teachers are relatively free of anxiety and stress, then students can be expected to also be relatively free of anxiety and stress in the classroom setting. Stipek referred to this as students reflecting what they see in their teachers. Therefore, the frustration and anxiety experienced by teachers in high-stakes educational environments are likely to be felt by their students also.

Another way of explaining test anxiety is to consider it as an interactional or transactional process. Bandura said that test anxiety develops in a social context. He described reciprocal determinism as the constant interaction of factors that are personal characteristics, behaviors that happen in reaction to the behaviors of others, and behaviors that happen in reaction to situations. This interaction of people with different personalities, the different ways people react, and the different situations in which they find themselves cause what they think, feel, or do in reaction. Therefore, individual behaviors are seen as not being so much individual as being determined within, and in part by, the environment. In other words, one way to consider test anxiety might be to see it as an interactive process that takes place during the test situation. One must understand how personality, behavioral characteristics, situations, and backgrounds interact during a test situation to understand the problems that test anxiety causes for individuals in test performance. Behavior patterns that develop in the interactions that occur in a family play a role in the formation, growth, and continuation of an individual's anxiety. Likewise, behavior patterns that develop in school affect the development, growth, and continuation of test anxiety for students; and so, it follows that behavior patterns that develop in workplaces affect the perpetuation of test anxiety in adults.

People are also seen as having differing levels of ability to deal with stress or anxiety that results from specific situations. Coping, according to Lazarus and Folkman, is the effort involved in managing stress and stressful situations in avoidance of distress that could be caused by these situations. If a person's

coping mechanisms are faulty or nonexistent, the result is test anxiety. As mentioned earlier, emotions are not seen as having an impact on causing test anxiety. Rather it is worry that affects the cognitive, or thinking, processes. Providing immediate feedback about how one is performing helps to alleviate test anxiety. Coping strategies can be learned and used to help alleviate test anxiety. One example of a proven, successful coping strategy is the use of cognitive-behavioral strategies, or positive self-talk, as a successful way of coping with test anxiety. If an individual is not performing up to his or her potential, test anxiety can increase.

Concerns With the Current Practice of Standardized Testing

Test anxiety in its current educational context is an issue often linked to the standards movement. Many experts have voiced the concern that using standardized testing as the sole means of measuring academic progress in itself may be responsible for an increase of test anxiety over the course of the standards movement from the space race era of the 1960s to the present time under NCLB regulations. The literature concerning test anxiety from the past 25 years is reflective of this. Therefore, discussion of the problem of test anxiety involves discussion of the standards movement, and discussion of the standards movement involves discussion of test anxiety.

Researchers originally examined the test anxiety issue involving students largely from junior high through adulthood. Because elementary students were not exposed to standardized testing from the beginning of the standards movement, elementary test anxiety did not become a serious focus of concern in the United States until after the implementation of NCLB. As early as 1960, Sarason and colleagues described the dependent connection of elementary students to their teachers and the tendency of children to reflect the outlook and attitude of their teacher or other influential adults in their lives. It even affects their attitudes toward themselves. These connections between the students and teacher resulted in test anxiety in elementary students becoming an area of special concern. Standardized testing that determines not only student success but also school success is now done in all elementary schools. Sarason pointed out that researchers agreed that children begin developing test anxiety in reaction to evaluative situations even before entering public schools. Conducting research has been somewhat problematic because elementary students tend to react even to the presence of a researcher in their educational environment. Factors other than actual test or practice test situations in the school experience can influence school stress (i.e., teaching methodology, instructional pacing, school climate). More recent literature on test anxiety has largely focused on negative concerns over test anxiety that have arisen since the beginning of the standards movement.

Test anxiety and its assessment have been studied extensively in Western culture. Theories about test anxiety, as reviewed by Bodas and Ollendick, included cognitive interference; low performance ability and poor study habits; disabilities in processing and organizing information combined with difficulty in remembering or recalling information; unrealistic parental expectations combined with school failure; emotionality (meaning bodily reactions such as an increase in heart rate or rapid breathing) versus worry (troubled or off-task thoughts); varying levels of physical, or physiological, reactions to anxiety that either facilitate or hinder test performance; and the tendency of test-anxious children to have an anxiety reaction to specific situations (or trait anxiety) in addition to state anxiety or comorbid disorders. Comorbid disorders means that multiple disorders are experienced at the same time. For example, a student with a learning disability or hyperactivity disorder might be expected to have test anxiety as well. Different assessment tools, including questionnaires (first developed for adults, and then children), measurement scales that consider different combinations of factors, self-report narratives, measurement of changes in bodily reactions, and observations of behavioral reactions, have been used to measure the extent to which test anxiety affects students and adults. Treatments have included relaxation therapy; systematic desensitization (creating an immunity to test anxiety by repeated exposure to the situation, much like decreasing an allergic reaction by exposing a patient to small doses of the allergen that causes the reaction); cognitive behavioral approaches (teaching one to use positive self-talk to convince oneself that one is prepared for the test and should not be nervous); test-taking strategies (including study skills and reading strategies); focus on the effect on individuals; and the effect according to socioeconomic status (the only contextual variable that has been considered). It is widely accepted as fact that students from low socioeconomic status can be expected to be less

successful on standardized tests than their more affluent counterparts. Students from a low socioeconomic status have also been shown to be more likely to experience test anxiety than their more affluent counterparts. At the present time, a great deal of energy and funding are being expended in the educational system in an attempt to counteract the negative effects of low socioeconomic status on test performance.

Test anxiety continues to be a concern in both the United States and internationally for preschool-age children, students, and adults. Preschool-age children are being taught test-taking skills and how to fill in bubble answer sheets in child care programs. Parents are keeping their 5-year-olds out of kindergarten to give them an extra year of test preparation before they enter competitive elementary schools. In 2006, according to Jervey, 70% of teens were reportedly cheating on tests to improve their chances in the competitive arena of gaining acceptance into desirable colleges. Internationally, governments are both mandating that testing be done and mandating moratoriums on testing young children. School systems in the United States are experiencing record teacher burnout and turnover as teachers flee the high-pressure accountability requirements of the field of public education for K–12 students. The acceptable passing level for standardized testing under NCLB continually increases until the 2013–2014 school year, at which time 100% of students will be expected to achieve passing scores on standardized tests measuring mastery of school curriculum, barring a change in policy.

Many schools are beginning to fall under close scrutiny both by the public and by the federal government. Schools that do not achieve the prescribed level of Adequate Yearly Progress (a legislatively defined increase in the pass rate of the school from year to year) face possible restructuring. What that restructuring means to schools, school systems, and students and their families has not yet been totally defined. Additionally, at the present time, a limited number of states are piloting an alternate method of certifying Adequate Yearly Progress that looks more closely at the amount of growth of individual students and subgroups of students (i.e., subgroups of differing ethnic backgrounds, or subgroups of children receiving special education services within the school). All of these factors play a role in influencing the test anxiety experienced by everyone involved in the educational system. Test anxiety today encompasses much more than political debates, athletic performances, or corporate promotions. Understanding the nature, the causes, and the factors that tend to occur in conjunction with test anxiety can help enable people to develop or strengthen much-needed coping strategies. Test anxiety can be viewed as a universal problem with varying degrees of magnitude and serious repercussions.

Penny L. Burge and Elizabeth V. Heath

See also Anxiety; Cognitive Behavior Modification; Motivation

Further Readings

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist*, *37*, 122–147.

Friedman, I. A., & Bendas-Jacob, O. (1997). Measuring perceived test anxiety in adolescents: A self-report scale. *Educational and Psychological Measurement*, *57*, 1035–1046.

Hembree, R. (1988). Correlates, causes, effects, and treatment of test anxiety. *Review of Educational Research*, *58*(1), 47–77.

Kohn, A. (2005). Unconditional teaching. *Educational Leadership*, *63*(1), 20–24.

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal and coping*. New York: Springer.

Liebert, R. M., & Morris, L. W. (1967). Cognitive and emotional components of test anxiety: A distinction and some initial data. *Psychological Reports*, *20*, 975–978.

McDonald, A. S. (2001). The prevalence and effects of test anxiety in school children. *Educational Psychology*, *21*(1), 89–101.

Sarason, I. G. (1984). Stress, anxiety, and cognitive interference: Reactions to tests. *Journal of Personality and Social Psychology*, *46*(4), 929–938.

Supon, V. (2004). Implementing strategies to assist test-anxious students. *Journal of Instructional Psychology*, *31*(4), 292–296.

TESTING

Testing is one method for collecting information to inform decision making. It involves eliciting individual responses to a set of structured tasks, often with the intent of obtaining a score that is representative of broader skills and/or knowledge. In most circumstances, test scores are used with information collected through other methods (e.g., record review, interview, observation) when making important decisions. In the following sections, a historical perspective on education testing is

provided, as well as information on testing purposes, testing types, test development, and standards for test development and use. A variety of current issues in educational testing are also described.

History of Testing and Education

Although examples of testing likely go back to the very beginning of humankind, some of the earliest accounts of tests being used to inform educational practice are from China during the time of the Han dynasty (circa 200 B.C.). Prior to the development of a testing system in China, men were selected for governmental positions based on their royal status. Following the development of an examination system, men from a variety of different backgrounds (e.g., low socioeconomic status) were selected for such positions based on successful test performance. These examinations are believed to have fostered a unified set of national values and educational goals across China. Following the development of these tests, educational programming was designed to promote the development of associated skills and knowledge among students, such that they would perform well on the tests in the future.

In the late 19th century, as the scientific method was becoming more widely used to study aspects of the human mind, strategies and tools were developed to systematically measure qualities of the human mind and its functioning. Wilhelm Wundt, who is often considered the first psychologist, developed methods for measuring sensory and perceptual experiences. His student, James Cattell, brought these early mental tests to the United States and then used them to test his university students. Around the same time, Alfred Binet co-created what has been considered the very first test of intelligence: the Binet–Simon scale. This test went through several revisions over time. Other tests of intelligence were also developed (e.g., Wechsler Scales, Woodcock–Johnson). These tests were originally used within schools to identify students who were considered to be in need of different educational programming.

During World War I, testing practices were used to identify individuals for military positions; draftees were administered the Army Alpha, which is considered to be one of the first intelligence tests administered in a large-scale fashion. Following World War I, school administrators began to increasingly administer similar brief tests of intelligence on a large-scale basis to all children within a given school. Tests similar to the Army Alpha were also administered to college applicants to help in making scholarship decisions. This was intended to allow for scholarship decisions to be less influenced by the quality of a student's past educational experience, and more by a student's potential for academic success.

In the past few decades, tests have become much more carefully developed, and results from standardized group and individual tests have become much more systematically used to inform a variety of decisions. Many laws and regulations now provide the foundation for the use of test results to inform decision making within educational settings. For instance, the No Child Left Behind Act of 2001 requires that schools and districts demonstrate that students are making adequate yearly progress, which is often determined based on student test performance. Furthermore, the Individuals with Disabilities Education Improvement Act of 2004 provides guidelines for how students are to be determined eligible to receive special education services. Tests are often used to help inform these eligibility decisions.

Despite the fact that tests have become increasingly used to inform decision making, criticism has also grown regarding the extent to which many standardized tests are equally effective in measuring student skills across all student groups. Several court cases have brought attention to the need for tests to be developed such that they are not biased in the measurement of skills among certain groups of students (e.g., English-language learners, students with disabilities).

Testing Purposes

Tests are administered within educational settings for a variety of different reasons. Some of these include prediction, screening, instructional planning, progress monitoring, eligibility for special programs, and accountability. When there are particularly important consequences attached to test score results, they are considered high-stakes tests. Examples of tests used for each of these purposes are described below.

Prediction

Some tests are developed in order to predict success in future educational settings. For instance,

college admissions tests such as the SAT, ACT, and GRE are given because of their predictive qualities; they are believed to be good indicators of how a student will perform in college or graduate school. Such scores may be used as part of a college or graduate school admissions process. It can be challenging to determine whether such tests are valid for the purpose of prediction, because students who do not perform well are often not selected for college admission, and therefore it is impossible to know how they would have performed had they been admitted.

Screening

Brief tests may be given on a large-scale basis (i.e., to all students) as part of a screening process in order to help in the identification of students who may be in need of special assistance. After screening, additional testing and other types of assessment data are needed before decisions about changes in educational programming are made for those students who initially perform poorly on the screening test. In many schools across the nation, hearing and vision screening occur for all students. Those who perform below a certain cut-point on such tests are then referred for additional testing to determine whether there is a serious problem with their vision or hearing. More recently, early literacy screening measures have been developed and used to identify students who are struggling in reading skill development. Following additional assessment, it may be determined that the student needs supplemental educational programming.

Instructional Planning

Some tests may be given to students prior to instruction in order to allow teachers to tailor instruction to meet identified needs. By learning through administration of a test what students know and do not yet know, a teacher can choose to emphasize certain content and de-emphasize other content that students already know and understand. This purpose is in contrast to what is perhaps the more common use of a test: that of determining whether students learned what was expected *after* the material has been taught. Ideally, testing for the purpose of instructional planning would occur often in order to ensure that instruction is targeted to student needs and therefore is as efficient as possible.

Progress Monitoring

A related purpose of some tests is to monitor progress toward long-term goals. Testing for progress monitoring can assist teachers with deciding when a change in instruction is needed to facilitate student learning. To be effective, progress monitoring tests must include knowledge and skill expected to be developed over a longer period of time, and they must be sensitive to small increments of growth in skill and knowledge over time such that they can be used on a frequent basis to determine if instructional changes are needed to better facilitate student learning.

Eligibility for Special Programming

Although brief tests used for screening purposes can be helpful to initially identify students who may be in need of special assistance, a more comprehensive, individualized assessment is warranted prior to considering a student eligible for longer-term special programming, such as special education. Longer and more reliable tests may be administered to determine whether students meet criteria for being in need of long-term special programming. However, whenever such important decisions are made, it is important that data are collected using a variety of assessment methods.

Accountability

Tests are increasingly being used and further developed to hold schools accountable for student outcomes. Such tests are intended to be administered to all students, and they serve as an indicator of whether resources are being used efficiently to promote student learning in schools. Results are publicly reported to members of the community and may be used to apply sanctions or awards to certain schools where students make limited or exceptional progress.

Test Development

Depending on the purpose for which it is intended to be used, a systematic process might be applied in the development of a test. Some tests, such as those that teachers might use to inform instructional planning, are often developed rather quickly. Other tests, particularly those that are intended to be used to make high-stakes decisions and those that are intended to allow for comparison of student performance to a

large representative sample, may involve more systematic procedures for test development. Some common aspects of more systematic test development are described below and include construct conceptualization, item development and selection, pilot testing, standardization, reliability studies, and validity studies.

Construct Conceptualization

In the initial stages of test development, test authors review research in the area in which student skills and knowledge are intended to be measured. The area intended to be measured is often referred to as the "test construct." Prevalent theories and conceptualizations of the given construct are used to provide a framework for how the test is organized. For instance, the Cattell-Horn-Carroll theory of intelligence (CHC theory) was used to develop the Woodcock–Johnson Test of Cognitive Abilities, third edition. Subtest tasks were constructed to measure the various factors according to this theory of intelligence.

Item Development

Once a framework for various subareas to be tested is developed, authors develop tasks and/or items to measure student performance in the given subareas and the overall construct. For norm-referenced tests, item performance is intended to help in distinguishing between students at a variety of different skill levels, and therefore both easy and difficult items are developed. For criterion-referenced tests, item performance is intended to help in distinguishing whether examinees have attained a certain predetermined level of competence, and so many items may be developed at a particular criterion level of difficulty. Items may be created to allow examinees to simply select a correct response from a set of possible options (e.g., multiple-choice format), or they may require the examinee to construct an answer or perform some other action. As test authors seek to measure a particular construct, they attempt to design items that will not require skills in addition to those that are intended to be measured. For instance, paper-and-pencil tests may require that students have adequate writing skills, which may be skills not related to what the test is intended to measure, if it is intended to measure social studies knowledge. In this case, a better testing method might be to have students respond orally.

Pilot Testing

Following item development, a sample of examinees might be recruited to test out the items. Those involved in pilot testing may be asked to talk out loud as they answer questions to ensure that they are engaging in the thought and performance patterns that are expected if the test is measuring the intended construct. Results from the pilot testing may help the authors decide whether certain items that were developed are appropriate for inclusion. If some items are found to have poor discrimination (i.e., not clearly distinguish between students of different skill and ability levels), they may be eliminated from the test. Similarly, items might be selected to appropriately measure student skills at a variety of different levels, or to measure well at one particular skill level. Items might also be analyzed by sensitivity and bias review panels to determine the extent to which they are sensitive to aspects of different cultures that may be present in the target population, and the degree to which items may be biased against students of particular backgrounds (e.g., gender, race, socioeconomic status).

Standardization

If the test is intended to allow an examinee's performance to be compared to that of a larger population of students, the test developers may seek to have the test administered to a group of students selected to represent characteristics of that population. Careful procedures are used to ensure that an adequate number of students from each age, gender, socioeconomic status, racial, ethnic, and geographic region group are tested and included in the standardization sample, such that they accurately represent characteristics of the population to which test performance is intended to be compared. Specific administration rules are presented in examiner testing manuals to ensure that the test is administered the same way to all of these individuals, to allow for appropriate comparison of skills. Performance of these students is summarized and consequently used as a comparison group when calculating scores of examinees who take the test later in time.

Evidence of Reliability

Test authors typically collect information to demonstrate a high level of test reliability. Reliability means that the test measures similarly across different

times, scorers, and items within the test. To demonstrate reliability, test authors may test the same group of individuals at two different times and correlate the test scores to determine if the test measures similarly (i.e., test-retest reliability); they may calculate a percentage of agreement in how two examiners score a given examinee's responses; and they may calculate coefficient alphas, which can serve as a measure of internal consistency within a test.

Evidence of Validity

A variety of different studies may be conducted to investigate the validity of a test. Factor analytic studies using examinee performance data can help identify whether performance patterns follow those expected based on the theory from which the test was derived. Examinee performance might be compared to performance on tests or other assessment tools that are considered to measure similar constructs. When such comparisons are done at the same time, this can provide evidence for concurrent validity; when comparisons are made such that the target test is used to predict future performance, this can provide evidence for predictive validity. Evidence of validity may also be provided by examining the extent to which examinee performance patterns follow those anticipated; for instance, that students with particular clinical diagnoses perform consistently in an anticipated direction (e.g., students with dyslexia perform far below average on a test of reading skills). As standards for testing have changed, the types of validity evidence that test authors are expected to provide has also changed. In the next section, information on standards for testing is provided.

Standards for Testing

In 1999, a joint committee of individuals selected to represent the American Psychological Association (APA), the American Educational Research Association (AERA), and the National Council of Measurement in Education (NCME) published a revised set of testing standards. According to these standards, the responsibility for appropriate testing is allocated to both test authors and test users. The developed document of test standards consequently delineates expectations for both test development and test use. The standards address expectations for (a) test construction, evaluation, and documentation; (b) fairness in testing; and (c) testing applications. Changing notions

of validity brought about changes in evidence needed to meet the new standards. More specifically, test developers are expected to document evidence for the appropriateness of test content, response processes, internal structure, relation to other variables, and consequences of testing. Special attention is given within the standards for testing students of diverse language backgrounds and those with disabilities. Test users are expected to follow standards for reviewing, selecting, administering, and interpreting tests that are outlined in the published document. In order to further guide application of the standards by test developers and users, the *Code of Fair Testing Practices in Education* was developed to be consistent with the new standards.

Types of Tests

Tests differ in a variety of ways. Some common differences in test uses that influence how various tests are developed and how they should be used are described below.

Criterion- Versus Norm-Referenced Tests

Criterion-referenced tests (CRTs) are developed and intended to measure whether an examinee has met a predetermined level of skill and knowledge. Norm-referenced tests (NRTs) are developed and intended to allow for comparing examinee performance to that of other, similar individuals. Criterion-referenced tests tend to have items that are of similar difficulty and that represent skills that are expected among students who meet the minimum criteria. Norm-referenced tests typically are developed to include items that measure performance across a wide range of student skill levels.

Group- Versus Individually Administered Tests

Group-administered tests are those that can be administered in a classroom setting, with many students participating in the test administration at a single point in time, and a proctor leading test administration. Individually administered tests are given in a one-on-one setting, and tend to allow the examiner to more carefully determine whether the test can be considered an adequate measure of the students' skills at a given point in time, based on examinee behavior. Group-administered tests often require that the examinee has

a particular set of skills (e.g., can read test items, follow directions independently) in order for the test to allow for accurate measurement; individually administered tests tend to allow the examiner more control over the testing environment as well as over the questions administered. This can help the examiner determine whether the test is effectively measuring the intended construct for a given student. For related reasons, group-administered tests tend to be less reliable than individually administered tests.

Recent Issues in Testing

Several developments over the past few decades that have influenced testing are described in the following sections.

Computerized Testing

Computers are increasingly being used in test administration and scoring. There are many benefits to this method of test administration. It can allow for efficient scoring and the provision of feedback quickly to teachers to inform their future practices. With the advent of computer adaptive testing, in which items are presented to students based on their earlier performance on the test, examinee performance level can be determined accurately with fewer items. However, there are also many challenges with this mode of test administration. For instance, in order for results to be fair, students may need similar levels of proficiency in completing work on a computer. Those who are not as familiar with working on a computer may be at a disadvantage. Furthermore, test security must be carefully monitored, particularly if testing is conducted over the Internet.

Bias in Testing

One common criticism of testing is that test items are often developed from the majority culture, and therefore may include item content that is not as accessible or familiar to those from other cultures. When students from different gender, racial, ethnic, and socioeconomic subgroups tend to perform differently on a test, it is questionable whether the test is measuring appropriately across all student groups. It is important to recognize that subgroup differences do not necessarily denote bias; more investigation is

needed to determine whether bias is present. Both expert review of items and statistical analytic methods (e.g., differential item functioning, or DIF) can be used to investigate the extent to which a test is biased for a particular subgroup of students. If many items display large DIF, it may indicate that the test is biased for a particular group of students.

Test Adaptations

Although tests are often developed with very strict administration procedures in order to allow for measurement of intended skills, and to allow for comparison of skills across those within a standardization sample, it can sometimes be the case that individual students have particular characteristics that make it difficult for them to demonstrate their skills under standardized conditions. For instance, a student with a reading disability may need to have math test items read aloud in order to demonstrate underlying math skills and knowledge. Such a change in standardized administration procedures is often referred to as a test accommodation. According to federal law, such changes must be provided to students with disabilities when appropriate. It can often be difficult to determine whether an accommodation is appropriate, and test developers and users are encouraged to have evidence to show that the change is appropriate and allows for appropriate measurement of the intended skills.

With the growing linguistic diversity present in our nation today, tests are being increasingly adapted in ways to allow students of different languages to participate. A variety of measurement challenges exist when translating test items into a different language. For instance, words may not have the same meaning across the different languages. Furthermore, if the student is being educated in English, he or she may know academic words in English better than in the native language. Determining how to test students of different language backgrounds can be very difficult, and individuals are encouraged to consult the test standards document for guidelines on how to test such students.

Universal Design for Assessment

With the growing diversity of students with various needs, test developers are being encouraged to

consider more carefully the wide variety of target students when they initially design and develop tests, such that later accommodation and adaptation are not needed. Using the principles of universal design, it is anticipated that tests can be developed such that they measure more accurately the intended skills and knowledge across all students. Some of these universal design characteristics include the following: precisely defined constructs; accessible, nonbiased items; amendable to accommodations; simple and clear instructions and directions; and maximum readability, comprehensibility, and legibility.

Sara Bolt

See also High-Stakes Testing; Reliability; Standardized Tests; Validity

Further Readings

AERA, APA, & NCME. (1999). *Standards for educational and psychological testing:* Washington, DC: American Educational Research Association.

Code of Fair Testing Practices in Education. (2004). Washington, DC: Joint Committee on Testing Practices.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory.* New York: Harcourt Brace Jovanovich.

Lyman, H. (1997). *Test scores and what they mean* (6th ed.). Boston: Allyn & Bacon.

Kubiszyn, T., & Borich, G. D. (2006). *Educational testing and measurement: Classroom application and practice.* Hoboken, NJ: Wiley.

Maddox, T. (Ed.). (2003). *Tests: A comprehensive reference for assessments in psychology, education, and business* (5th ed.). Austin, TX: Pro-Ed.

Murphy, L. L., Spies, R. A., & Plake, B. S. (Eds.). (2006). *Tests in print VII.* Lincoln, NE: Buros Institute of Mental Measurements.

Salvia, J., Ysseldyke, J., & Bolt, S. (2007). *Assessment in special and inclusive education* (10th ed.). Boston: Houghton-Mifflin.

Spies, R. A., & Plake, B. S. (Eds.). (2005). *The sixteenth mental measurements yearbook.* Lincoln, NE: Buros Institute of Mental Measurements.

Thompson, S. J., Johnstone, C. J., & Thurlow, M. L. (2002). *Universal design applied to large scale assessments* (Synthesis Report 44). Minneapolis: University of Minnesota, National Center on Educational Outcomes. Retrieved November 13, 2006, from http://education.umn.edu/NCEO/OnlinePubs/Synthesis44.html

THEORY OF MIND

Theory of mind is the ability to understand the thoughts and feelings of other people and oneself. It is generally used to refer to a stage-like development at the age of 4 years, when children start to understand *belief.* Until this age, children are good at predicting aspects of behavior. After this age, they become increasingly good at predicting other people's mental states. It is referred to as a theory because we understand people's thoughts and behaviors in terms of a set of mental state concepts, concerning beliefs and desires, and a set of rules relating these concepts to behavior and experience. For example, if I *want* a cookie and I *think* there is one in the cookie jar, then I will go to the cookie jar. For emotions, we may use a less theory-like method, imagining ourselves in someone else's place and asking, "If I were in this situation, how would I feel?" This entry describes the development of children's understanding of belief, related developments, precursors, and the role of theory of mind in characterizing autism.

Understanding Belief

Even very young children know that people behave in ways that satisfy their desires: Two-year-olds are quite able to predict that if someone wants a cookie, he or she will go to the jar containing cookies. However, children do not properly understand belief until about 4 years of age. Beliefs generally can be disregarded in everyday interactions because they are normally true. As a result, others' behavior can be predicted fairly accurately by reference to the actual state of the world. This is not always the case, however, particularly when beliefs are false. In this case, it is necessary to refer to the others' mental states.

The classic test of false belief understanding involves the following short scenario, with dolls and props:

Maxi puts his chocolate in Location 1, then goes out to play. While he's away, his mother moves the chocolate to Location 2. A bit later, Maxi comes back. Where will he look first for his chocolate?

Three-year-olds typically say he will look in Location 2: They predict his behavior on the basis of what is really the case. By the time they are 4 or 5 years old, most children realize that Maxi will look in Location 1, where he *thinks* the chocolate is; they now realize that a person's behavior is based on how the person represents the world.

Another version of the false belief task is the unexpected contents task. Here, a child is shown (for example) a Smarties container and is asked "What's in here?" After answering "Smarties!" the child is shown that it contains only a pencil. The container is closed again, and the child is asked, "What did you think was in here before I took the top off?" Three-year-olds typically answer "a pencil," despite having just had a false belief to the contrary. They also predict that their friend, who has not seen inside yet, will think there is a pencil inside. This suggests that people may not have special insight into their own mental states; at least for outdated beliefs, people reconstruct them in a theory-like process, as they do when predicting others' mental states.

Understanding belief opens up a whole new world of possibilities for children. Studies have shown that once children understand belief, they start to selectively withhold information from others. They are able to keep secrets (sometimes), and play games like hide-and-seek properly. Younger children enjoy the rituals of hide-and-seek immensely, but do not seem to understand that the seeker should not know where the hider is. A less positive consequence of understanding belief is that children begin to lie. Three-year-olds may deny things, but this is probably a learned formula, used to ward off embarrassing questions. After the age of 4 years, however, children realize that it is possible to plant false beliefs in others' minds, and they become increasingly adept at this.

Another advantage of understanding mental states may be increasing self-control. Recognizing that you have a tendency to behave in a certain way is the first step toward modifying that behavior. A variety of self-control or "executive function" tasks are used in theory of mind research. These include a version of the popular game "Simon says," in which children have to perform actions only if the instruction is preceded by the words "Simon says." So, after "Simon says, 'put your hands on your head,'" you should do so, but after "Put your hands in the air," you should not. Children become able to inhibit actions appropriately

in situations like these at the same time that they pass false belief tasks, and the two sets of abilities are strongly related. The direction of causality is controversial; theory of mind plausibly allows increased self-control, but children may also need some ability to control their own thought processes in order to adequately reflect on mental states.

Earlier and Later Understanding

Most research has focused on false belief because it marks a watershed in our understanding of mental states. Nevertheless, although younger children cannot make explicit judgments about belief, they have some implicit sensitivity. For example, 3-year-olds will look to the empty location when anticipating Maxi's return in standard false belief tasks, yet answer the explicit question incorrectly. Recently, one study has found that even 15-month-old infants show sensitivity to false belief through their looking behavior. The nature of this implicit sensitivity is an area of intense interest and controversy; researchers are divided over whether it indicates innate theory of mind ability or some simpler form of social cognition.

Although young children cannot explicitly ascribe beliefs, they show other forms of social understanding from infancy. For example, by 1 year, infants can follow other people's gaze direction. This ability is extremely useful in orienting the child's own attention to interesting things in his or her environment; furthermore, it allows children to take into account the speaker's focus of attention to infer the meaning of a new word. Gaze understanding during infancy seems to be implicit: Children can only explicitly judge what someone is attending to from the age of roughly 3 years.

Children also begin to use pretense in their play at around 18 months. Pretend play is closely related to theory of mind; pretense is a mental state very much like belief, but can be understood in simpler terms. Children's early pretense is usually social, often encouraged and assisted by a parent, and they can follow events in shared pretend worlds. For example, if, during a pretend tea party, the mother accidentally knocks over a cup, 2-year-olds realize that pretend tea has been spilled everywhere and interpret the mother's wiping-up behavior appropriately. In other words, children interpret another's behavior as directed toward a false situation, just as in the false belief task: Maxi's behavior must be predicted according to his false

belief about the situation. However, there is a critical difference. In pretense, typically everyone is aware that the pretend situation is not true. In false belief, in contrast to the child, Maxi is unaware that the belief situation is false. Maxi genuinely wants to retrieve his chocolate, rather than indulge in pretend play, so the puzzle is why he would base his behavior on an untrue situation. It is not until around the age of 4 years that children realize that people may evaluate the same situation differently. However, this realization is unnecessary for pretense, so for practical purposes, children can understand pretense earlier than they understand false belief.

Considerable development takes place after 4 years. Initially, children do not understand that knowledge can be partial. For example, 4-year-olds fail to understand that they cannot tell the weight of an object by simply seeing it. Similarly, they judge that someone who sees a small, unidentifiable portion of an object will know what the object is, even though children themselves could not identify it when they first saw the small portion. Until around the age of 5 years, children think that any perceptual access to an object is sufficient to know all about it.

Even older children have persistent problems with mental states that refer to other mental states. For example, 5-year-old children think that if one asks Maxi (who has a false belief), "Do you know where your chocolate is?" that Maxi will say "no." They fail to realize that Maxi believes his own beliefs to be true. Children have similar difficulties with belief-based emotions. Even 2-year-olds have a basic appreciation of emotions in terms of the kind of circumstances that lead to happiness, sadness, anger, and so on. However, consider the following situation: A person likes cola and hates milk. We have played a mean trick on her by giving her a cola can full of milk. Five-year-olds will correctly judge that on first seeing the can, she will think it contains cola, but they will also judge that when she first sees it, she will be unhappy (because she does not like milk). They fail to predict her emotion on the basis of her belief, predicting it instead according to objective reality. By 6 or 7 years of age, children can predict beliefs about beliefs, emotions based on beliefs, and so on. The foundations of adult understanding are laid: Throughout our lives, we become increasingly able to infer the thoughts of others. The pinnacle of theory of mind ability is probably seen in a good novelist or actor.

Influences on Development

What drives the development of theory of mind is unknown. Development may depend partly on brain maturation; theory of mind development follows a similar timetable across a wide range of cultures, from the United States, to Japan, to the central African rain forest, in spite of very variable environments and child-rearing practices. One possibility is that there is an innate "theory of mind module," an evolved brain mechanism specialized for theory of mind reasoning. Another possibility is that development depends on the general maturation of the frontal lobes, the part of the brain associated with executive function.

It is also clear that experience plays a major role in development. The quality of the early maternal bond predicts later theory of mind ability. This reflects the mother's tendency to think of her child as a thinking, autonomous individual, a trait known as *maternal mind-mindedness*. This affects how the mother interacts with the child. If child-directed speech includes the appropriate use of terms like *want, think*, and *know*, children's mental state understanding is enhanced. This probably interacts with family background. The presence of older siblings has been shown to promote theory of mind development. Older siblings have superior theory of mind and may "tutor" children; they also provide a ready partner for pretend play, which also encourages theory of mind development. The presence of a large extended family has also been found to be effective.

Developmental Difficulties

People with autism appear to lack a theory of mind. Autism is a lifelong developmental disorder characterized by difficulties in social interaction, difficulties in communication, and a narrow range of interests and repetitive activities. It is a complex condition that has resisted simple explanation. It probably has multiple psychological causes and consequences, one of which is an impaired theory of mind. People with autism perform poorly on tests of theory of mind and are baffled by social interaction. When communicating, they tend to speak too loudly, too quietly, or too quickly, and repeat themselves and others, apparently failing to appreciate the informational needs of the listener. They tend to be overly literal and do not make the

inferences necessary to go from what was said to what was meant; they are unable to infer the thoughts and intentions behind the words. A lack of theory of mind is not a full explanation or description of autism, but it is a major symptom.

Martin John Doherty

See also Autism Spectrum Disorders; Social Development

Further Readings

Carpendale, J., & Lewis, C. (2006). *How children develop social understanding.* Oxford, UK: Blackwell.

Doherty, M. J. (2008). *Theory of mind: How children understand others' thoughts and feelings.* Hove, UK: Psychology Press.

Frith, U. (2003). *Autism: Explaining the enigma.* Oxford, UK: Blackwell.

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science, 308,* 255–258.

Perner, J., & Ruffman, T. K. (2005). Infants' insight into the mind: How deep? *Science, 308,* 214–216.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development, 72,* 655–684.

TIME-OUT

Time-out from reinforcement, also referred to as time-out (TO), is a procedure in which positive reinforcement is not available for a specified period of time. During TO, the individual does not have access to positive reinforcers that are normally available in the setting. TO is a punishment procedure developed from laboratory animal research that evaluated the impact of temporarily suspending the opportunity to earn food reinforcement (i.e., time-out from reinforcement). Extrapolations of the laboratory procedure have been effective in changing behavior in many applied settings (e.g., classrooms; day care facilities; the home; institutions for children, adolescents, and adults with mental retardation or psychiatric disorders).

TO requires that sources of reinforcement (e.g., attention from the teacher, participation in an activity) are withdrawn for a brief period of time following the behavior that one wants to decrease. In the usual application, a location is selected that removes the child from the situation—in the classroom, this would be a special chair or corner of the room where the child must remain for the duration. However, the definition permits many other variations. If attention is the reinforcer, keeping the child in the situation but not attending to or calling on him or her for a brief period constitutes TO. In an effective and creative variation, children in a special education class received social reinforcement (praise and smiles) for their work. Each child wore a ribbon around his or her neck that signified that the teacher could provide reinforcement as the children worked. Any disruptive behavior was followed by TO, which consisted of removing the ribbon for three minutes. During this period, the child could not receive reinforcers normally administered. Other variants of TO have included a period in which TV or music is turned off or in which others are asked to leave the room to go to another activity while the person remains behind.

The effectiveness of TO depends on several conditions:

- TO ought to be administered immediately after the behavior.
- The duration should be brief (e.g., 1 to 10 or 15 minutes). More time is not invariably more effective.
- The behavior and TO procedure ought to be described to the child in advance so that all is explicit.
- TO should be preceded by a verbal warning that TO will be administered if a behavior reaches a certain level.
- When administered, the person should be told calmly that he or she has to go to TO, told exactly what was done, and how much time will be spent.
- Physical coercion (e.g., grabbing by the arm, forcing) should not be used to take the child to TO; this is likely to foster oppositional behavior and aggression.
- While in TO, no attention should be provided to the child.
- If the child refuses to go to TO, it is helpful to have a back-up aversive consequence (loss of privilege) that can be used instead.
- When the child is in TO, there may be some requirements—such as being calm for all or some of the time interval before returning to the environment.
- A timer or the equivalent should be used to track the duration so that both the person who administers the procedure and the child who is in TO have an objective way of tracking the time.
- Depending on the setting and location of TO, it is useful to be able to check on the child during TO to ensure he or she is safe.
- When the individual is not in TO, there ought to be a reinforcement program to foster prosocial behavior and behavior that is opposite of the punished

behavior. For example, if disruptive behavior in the classroom is the basis for invoking a 5-minute TO, there ought to be an explicit reinforcement (e.g., praise, tokens) program for academic performance and working on assignments.

Of all the guidelines, the most important one is the one most commonly overlooked. The effectiveness of TO depends heavily on the ongoing reinforcement in the environment. The "time-in period" (i.e., what is going on when individuals are not in TO) is critical. If the environment is highly reinforcing (e.g., teacher praise, engaging activity), removal from the reinforcement is much more effective than if this is merely isolation in a relatively sterile environment. Punishment does not teach what to do; it only focuses on what not to do, and the desired behavior does not automatically increase or occur as a result of punishment, whether TO or some other form. Effective application combines reinforcement for prosocial behavior with TO for the behavior to be suppressed.

The extensive use of TO, the availability of many variations, and recommendations in many parenting books and Web sites have contributed to misconceptions about the procedure and how it is to be used. Among the more common misinterpretations is that TO is a period to have the child think about or understand, to cool down, to feel remorse, and so on. All of these can occur, but they are not central to the procedure, something that might be inferred from research across multiple species. Also, longer TO periods are sometimes used, but only a few minutes are needed. A goal in classroom use is to return the child to the regular class as soon as possible so that instruction, teacher reinforcement, and peer interactions are not withheld longer than necessary. TO has been used as a justification for detention, expulsion, and extended periods of confinement (e.g., child locked in his room for hours or days). These are not TO and reflect misconceptions of the principles and technique on which TO is based.

Overall TO is an excellent alternative to many other procedures, especially those involving corporal punishment, shouting, or excessive reprimands. Punishment often has undesirable side effects (i.e., features that emerge that have no bearing on whether the procedure is effective). Among the familiar examples are emotional reactions (crying, shouting); aggression (hitting others, including the agent who administers punishment); and relationship issues (e.g., efforts to escape or avoid the punishing agent). TO is not immune from these side effects, but they are milder than with corporal punishment. Also, side effects can be attenuated further when TO is administered with positive reinforcement for prosocial behavior. TO should always be viewed as a supplement to a positive reinforcement program and is very effective in that context.

Alan E. Kazdin

See also Applied Behavior Analysis; Behavior Modification; Discipline; Parenting Styles; Teaching Strategies

Further Readings

Ferster, C. B. (1957). Withdrawal of positive reinforcement as punishment. *Science*, *126*, 509.

Ferster, C. B. (1958). Control of behavior of chimpanzees by time out from positive reinforcement. *Psychological Monographs: General and Applied*, *72*(8, Whole No. 461), 1–38.

Foxx, R. M., & Shapiro, S. T. (1978). The time out ribbon: A nonexclusionary time out procedure. *Journal of Applied Behavior Analysis*, *11*, 125–136.

Hall, R. V., & Hall, M. L. (1998). *How to use time out* (2nd ed.). Austin, TX: Pro-Ed.

Kazdin, A. E. (2001). *Behavior modification in applied settings* (6th ed.). Belmont, CA: Wadsworth.

TOKEN REINFORCEMENT PROGRAMS

First developed and implemented in the 1960s by pioneers such as Nathan Azrin and Gordon Paul, token reinforcement programs, or token economy systems, are behavioral modification systems that reward individuals with tokens for the completion of target behaviors. Tokens are exchanged later for backup reinforcers. To create a successful token economy, tokens, behaviors, reinforcers, and the ratio of exchange should be identified and understood.

An initial step in creating a token economy is choosing what materials or objects will be used as tokens. Tokens used in a token economy should not be inherently reinforcing, but instead should become reinforcing when paired with a backup reinforcer. Items such as baseball cards or stickers should not be used as tokens. Examples of effective tokens include tickets, holes punched in a card, washers, or checkers. Tokens should be durable, easy to carry, and easily accessible to the instructor. Most importantly, tokens

should be given immediately following the individual's execution of a target behavior.

When defining the target behavior, it is important to choose behaviors that are specific, observable, and measurable. Behaviors should be individualized for each student and should be tasks that the student has the basic skills necessary to perform. In the beginning of the program, the criteria for earning a token should be relatively easy (i.e., easier behaviors or fewer occurrences of behavior required), but should gradually become more challenging and require more occurrences of the behavior to earn a token.

Another element central to the token economy is the backup reinforcer. Multiple reinforcers should be offered in exchange for tokens at pre-established criteria. At first, backup reinforcers should be available for token exchange often, and should gradually become less available and harder to earn. Examples of possible reinforcers include items or activities that are naturally occurring in the classroom, such as a favorite classroom job or time to play with a classroom computer game. If a naturally occurring reinforcement is not available or is not reinforcing to the individual, then other options should be introduced into the learning environment, such as Silly Putty or candy. In the school setting, ethical standards dictate that essentials such as meals, communication, clothing, or medical care cannot be used as reinforcement.

The ratio of exchange is the number of tokens that will be required to purchase various reinforcers. This ratio is simply the number of tokens versus the price of items. At first, the ratio should be small, with the price of items gradually increased and the value of tokens gradually decreased. The reinforcers should also be changed frequently. Additionally, the price of essential items should be increased while keeping the price of nonessential items low, thus maintaining the motivation of the learner to earn additional tokens. This prevents satiation and habituation.

The token economy system is one of the most common classroom management systems. Teachers will often provide points or tokens to be exchanged later for rewards. Before introducing the token economy, teachers define the classroom expectations and rules, such as have quiet hands and feet, stay in the area, and listen to the teacher. After children understand the rules of the classroom, the teacher can add earning tokens for following the rules. Children can then redeem the tokens at the treasure chest or classroom store. The ultimate goal should be removal of the rewards and

token system. This is accomplished by gradually fading the frequencies of earning tokens and redeeming them at the store. In addition, the amount of tokens required to earn a reward can be increased to make it more difficult to earn the reinforcement. To ensure successful withdrawal, the tokens and rewards should be paired with verbal praise, and the amount of naturally occurring reinforcers should increase while the amount of secondary reinforcers should decrease.

The advantages of token reinforcement programs are numerous. First, they are easy to implement and are effective. Second, tokens serve as a transition between the target behavior and the reward. For example, the teacher can give a token in the hallway for remaining quiet while continuing to move the class through the transition. Third, token reinforcement programs can be implemented at the individual or classroom level and can be based on individual or group goals. In addition, they are often the foundation for more intrusive, complex behavior management interventions such as response cost, self-management, and behavioral contracting.

Finally, token reinforcement programs have been found to be an evidence-based intervention for children with behavior problems. For example, Kara McGoey and George DuPaul implemented a token economy and response cost system with four preschool children with attention deficit hyperactivity disorder. Participants earned buttons on a chart for following the rules of the activity and classroom. The children then earned a large button and stickers to take home based on predetermined criteria of small buttons earned. The token economy system decreased the disruptive behavior of the participants and was found to be acceptable by the teachers. Additional research has also examined the token economy system with school-age children and adolescents.

Kara E. McGoey and Kristin Rezzetano

See also Applied Behavior Analysis; Behavior Modification; Reinforcement

Further Readings

Cooper, J. O. (1987). Token economy. In J. O. Cooper, T. E. Heron, & W. L. Heward (Eds.), *Applied behavior analysis* (pp. 486–498). New York: Macmillan.
DuPaul, G. J., & Eckert, T. L. (1997). The effects of school-based interventions for attention deficit hyperactivity disorder: A meta-analysis. *School Psychology Review*, *26*, 5–27.

Liberman, R. P. (2000). Images in psychiatry: The token economy. *American Journal of Psychiatry*, *157*(9), 1398.

Martin, G., & Pear, J. (2003). *Behavior modification: What it is and how to do it* (7th ed.). Upper Saddle River, NJ: Prentice Hall.

McGoey, K. E., & DuPaul, G. J. (2000). Token reinforcement and response cost procedures: Reducing the disruptive behavior of preschool children with attention-deficit/hyperactivity disorder. *School Psychology Quarterly*, *15*(3), 330–343.

Pelham, W. E., Wheeler, T., & Chronis, A. (1998). Empirically supported psychosocial treatments for attention deficit hyperactivity disorder. *Journal of Clinical Child Psychology*, *27*, 156–167.

Sheridan, S. M. (1995). *The tough kid social skills book.* Longmont, CO: Sopris West.

Stage, S. A., & Quiroz, D. R. (1997). A meta-analysis of interventions to decrease disruptive behavior in public education settings. *School Psychology Review*, *26*, 333–368.

TRACKING

Tracking refers to the practice of grouping students according to achievement levels, either between or within classrooms, for the purposes of instruction. The term *ability grouping* is frequently used in place of tracking, especially when discussing within-class tracking in elementary school, and British researchers often use the term *streaming* in place of tracking. The official rationale behind the practice of tracking is that by grouping students of similar achievement for instruction, classroom instruction will be more appropriately tailored to students' needs, and both high- and low-track students will experience more rapid achievement growth. However, research shows that, in reality, tracking tends to increase educational inequality, with low-track students learning less and high-track students learning more than students in regular (middle) or untracked classrooms. This entry summarizes research that helps to explain the contribution of tracking to educational inequality. Tracking provides a powerful institutional status marker that affects both teacher and student behavior, and consequently the quality of instruction in tracked classrooms.

Magnifying Initial Differences in Student Achievement

The early research on tracking was frequently concerned with whether or not the track placement process was meritocratic and with describing students' opportunity for upward mobility. Are track placements based on legitimate criteria such as achievement and effort, or do particular groups of students have an unfair advantage in securing placement in the college-prep track, whereas other groups of students are relegated to low-track classrooms? Once a student is in the low track, what opportunity does he or she have to move into a college-prep curriculum? As James Coleman articulated in the introduction to Alan Kerckhoff's landmark book *Diverging Pathways*, in order for structural differentiation of students to lead to growing educational inequality, there need not be any bias in the assignment of students to learning environments within or between schools. If learning environments differ in their *intensity*, then increasing inequality could result even if assignment and opportunities for mobility are completely based on student achievement levels. In other words, differences in the quality of instruction in tracked classrooms will lead to growing educational inequality no matter how students are assigned to those tracks. This observation is essential to understanding the role of tracking in producing educational inequality.

It turns out that track placements from year to year are, in fact, mostly meritocratic; they are determined in large part by the achievement level of students. For example, most of the difference in the track placements of students from different race/ethnic groups can be explained by differences in students' grades and test scores when track decisions are made. Students from advantaged family backgrounds appear to have somewhat of a nonmeritocratic advantage in securing high-track placements, but most of the total difference in track placements among students from advantaged and disadvantaged backgrounds is still due to levels of achievement. Nevertheless, tracking makes an important contribution to growing educational inequality as students of differing initial achievement progress through school. Among Black and White students, for example, initial differences in achievement are magnified as White students progress through school in intense, high-track learning environments, and Black students in less-intense, low-track classrooms. By the time Black and White students leave high school, the initial achievement gap has doubled.

Certainly, instruction in some low-track classrooms is of high quality and is as effective as instruction in many high-track classrooms. But on average, low-track classrooms represent significantly reduced

learning opportunities for students. One study by Adam Gamoran of the effect of track differences found that the difference in achievement growth in high school among tracked classrooms was greater than that between high-school dropouts and students who stayed in school. Tracking does little to affect the average level of achievement in a school as a whole, but it has a powerful effect on the distribution of achievement and contributes to enduring educational problems of minorities and those from disadvantaged backgrounds. The remainder of this entry summarizes research that illuminates why low-track classrooms are frequently less effective learning environments than high-track classrooms.

Students' Reactions to Tracked Learning Environments

Low-track students are frequently found to be inattentive, withdrawn, and disruptive. The lack of engagement during classroom instruction and more general disinterest in school activities can be traced to (a) the lack of incentives for high performance in low-track classrooms; (b) negative peer group processes that exacerbate misbehavior and disidentification with school; and, as discussed in the next section, (c) deteriorating relationships with their teachers.

Lack of Incentives

An obvious explanation for disengagement among low-track students, and one that is frequently articulated by teachers, is the lack of educational incentives for performance in low-track classrooms. It is easy for adults to forget, but especially younger students, that the material benefits of education, such as a steady job with fair pay, lie in the distant future and don't factor very heavily in their day-to-day attitudes toward schooling. In the absence of an immediate material payoff, teachers rely on grading systems to communicate expectations and foster motivation. Tracking systems undermine the utility of grades in low-track classrooms because they give students a clear status assignment at the start of the year, before any effort has been exerted. For college-bound students, the admissions process gives students a strong incentive to pursue high grades. But for low-track students, most of whom will enter the workforce after school, there are few institutional linkages between high school and the workforce that reinforce performance in school. It

simply does not make much difference whether a low-track student receives an A or a B, or even a C. Grade-weighting schemes, which give extra points for high-track coursework, further erode the incentive for low-track students to perform in school because the system is "rigged" against them.

The link between incentives and student behavior was well-illustrated in a study by Francis Schwartz. Schwartz observed both high- and low-track middle school students throughout the course of their day. As they moved from formal instruction in core academic subjects to gym, art, and other elective classes, Schwartz noticed an interesting pattern of behavior. High-track students were just as likely to misbehave as low-track students, but only during informal instructional time that did not affect their future educational trajectory. In contrast, high-track students were highly responsive to teacher directives and sanctions during formal class time, whereas low-track students misbehaved and failed to complete classroom tasks. The track system, by removing any incentive for performance among low-track students, strips the low-track teacher of most of his or her authority.

Negative Peer Group Effects

The differentiation-polarization theory of Andrew Hargreaves and Colin Lacey suggests that peer group interactions exacerbate antischool behavior among low-track students. Differentiation-polarization theory is an example of a more general social identity theory of behavior applied to status differences among tracked students. Because low-track students are labeled as low achieving by the school system, they need to look elsewhere for a positive self-image. As they develop alternatives to school achievement, such as accomplishments in athletics, working on cars, and being a sought-after date on the weekend, they develop group dynamics that support their chosen alternative methods of obtaining this positive self-image. These group dynamics include monitoring others' behavior and sanctioning students who show an interest in school. Developing alternative sources of success, and really believing in them, requires collective effort, and that effort is undermined by individuals who conform to the school's definition of success.

Students are much more likely to be friends with other students in the same track. Same-track friendships develop in part because of shared social experience, but also because of a shared *reaction* to school.

Over time, student attitudes in different tracks become polarized, with antischool attitudes being concentrated primarily among low-track students. Case studies by Hargreaves and Lacey, and later Stephan J. Ball, illustrated this process. The low-track students they observed did not have antischool attitudes entirely because of their low-track placements, and many low-track students had positive attitudes toward school, but these researchers concluded that tracking greatly polarized the differences in attitudes and behavior between high- and low-track students. Early, small-scale quantitative analyses also provide support for differentiation-polarization theory.

In 1982, Michael D. Wiatrowski and colleagues challenged differentiation-polarization theory, suggesting that perhaps the link between tracking and antischool attitudes was not causal, that low-track students did have antischool attitudes, but they existed prior to their track placements. In their quantitative analysis of 1,620 male students, they found no evidence that tracking contributed to delinquent behavior, once prior delinquency was accounted for. However, their analysis had two methodological flaws. First, 18.7% of the students left the study between their sophomore and senior years, when delinquent behavior was measured as an outcome. If the most delinquent youth dropped out of school or otherwise did not participate, Wiatrowski et al. might have missed track effects due to attrition bias. Second, they considered track effects very late in the schooling process. Presumably, by sophomore year, antischool attitudes are already well developed, and ceiling effects preclude any great increase in the disparity between high- and low-track students' attitudes.

The causal nature of the differentiation-polarization process was later confirmed by John Abraham, who examined pro- and antischool attitudes in a high school setting where track placement was only formalized in the students' second year. He found that there was a dramatic difference in school attachment between high- and low-track students, but that this difference appeared only after students became tracked. In perhaps the most rigorous treatment of this question to date, Mark Berends found small but consistent negative effects of tracking on college expectations, disciplinary problems, and engagement in the last 2 years of high school. Considering the body of research on differentiation-polarization theory, both qualitative and quantitative, Berends finds a clear link between low-track placements and negative peer group effects.

The robust conclusion of tracking research is that by the time students reach high school, the negative effects of low status assignments manifest themselves in the way students think and behave in school settings. Being in low-track classrooms disenfranchises students, and this in turn leads to a lower probability of moving into academic courses in future years and of pursuing further education after high school. A student's expectations and aspirations are closely related to track placement. The reduced aspirations of low-track students are clearly evident in dropout rates. Indeed, track placement itself exerts a strong independent effect on dropping out, stronger even than its effect on achievement alone.

Teachers' Reactions to Tracked Learning Environments

Teachers respond just as negatively to a low-track teaching assignment as students do to being in a low-track classroom. In Reba Page's study of teachers' reactions to tracked learning environments, low-track students are perceived overwhelmingly as "discipline problems," "untrustworthy," and "irritating." As Marilee K. Finley noted, although teachers are not in complete agreement about the desirability of teaching the highest-track classes, they are in agreement about the difficulties of instruction in low-track classrooms. Why do teachers respond so negatively to low-track students?

A Crisis of Authority

One explanation, consistent with a lack of incentives for student engagement and effort in low-track classrooms, is that teachers' reactions stem from a crisis of authority in low-track classrooms. As noted previously, teachers have little recourse to force students to be on task or well-behaved, and low-track status designations further undermine the utility of grades. As one teacher in Jeannie Oakes's study said of low-track students, "They don't like me in a position of authority, these children don't like anybody in authority." Of course, antagonistic reactions to authority figures might be expected if teachers and other school personnel consistently identify a student as low status.

Many of the low-track teachers Oakes interviewed believed that the only way to deal with their students was to be excessively punitive. Similarly, the low-track teachers Mary Metz observed responded to their

boisterous low-track students with somewhat stronger disciplinary sanctions than in high-track classrooms. If teachers respond to a disengaged student body by placing an excessive emphasis on discipline, this is likely to strengthen a reciprocal pattern of negative interactions between teachers and students. Beyond that, though, two approaches prevailed in low-track classrooms: a reduction in pressure to succeed in academics, and use of highly structured activities. Low-track teachers believed that highly structured activities like worksheets would help control students, who would quickly get out of hand if given the opportunity. The reduced academic pressure amounted to a bargain of sorts, where the teacher allowed students to be inattentive, or off-task, in the hopes of avoiding more serious rebellion. In formulating and carrying out lesson plans, low-track teachers choose methods that allow them to closely monitor and control student behavior.

A Lack of Cultural Coherence

Another explanation is that teachers, most of whom were successful as students and who have an appreciation for academic material, have difficulty identifying with low-track students' needs and perspectives. In a case study of a single teacher who taught both high- and low-track English classrooms, Samantha Caughlan and Sean Kelly describe a situation in which a teacher who is quite successful in her high-track classroom fails to provide similarly effective instruction in her low-track classroom. In explaining the disparity, Caughlan and Kelly link a lack of coherence in the low-track classroom, both within the curriculum and with the students' lives, with the teacher's failure to identify with the students' perspectives, interests, and needs. One expression of this is the assumptions that low-track teachers frequently make about their students' family backgrounds. Although low-track teachers want their students to be successful and happy as much as any teacher, they feel held back by their students' home environments, which they believe negatively affect their students' dispositions toward school. For many teachers, a home environment supportive of education is seen as a prerequisite for their own success in the classroom.

Perceptions of students' home environments may act as the basis for a self-fulfilling prophecy for teachers. From the perspective of the teachers Page interviewed, for example, a student's classroom success

hinged not so much on the inherent qualities of the student, but on that student's home environment. Unfortunately, some teachers let their reasoning serve as an excuse for poor outcomes in the classroom. Rather than adopt a particular instructional approach to meet their students' needs in the context of sometimes-difficult home situations, the teachers' instructional approach was tailored around student *limitations.* The basic format and style of instruction was not much different from that in a high-track class. Teachers lectured, had students recite answers, and discussed texts. But class sessions lacked a sense of purpose, were tedious, or frequently veered off-topic and became nonacademic.

The Cumulative Impact of Tracking on Teachers' Work Lives

Low-track classrooms are a difficult environment for teachers to succeed in both because students are perceived as inherently difficult to teach, and because the students are more likely to have problems with achievement motivation and negative alignments to school in general. Making matters worse, in many schools, teachers are tracked along with the students. The matching of teachers with the hierarchically structured courses for students is known simply as *teacher tracking.* Teachers with less education, experience, and motivation are more likely to be assigned to low-track classrooms. Thus, teacher tracking pairs students who are the most difficult to teach with teachers who, in some ways, are least equipped to be successful. It is not surprising, then, that low-track teachers have a much lower sense of efficacy—their own perception of their ability to succeed in the classroom—than high-track teachers. Consequently, low-track teachers are less satisfied with their work lives. Lower levels of efficacy and satisfaction that persist from one year to the next are likely to further reduce the likelihood that low-track teachers will expend extra energy to reach disengaged students.

The Quality of Instruction in Tracked Learning Environments

Several studies have systematically observed classrooms of differing track levels using comprehensive coding schemes designed to describe the range of classroom activities in which students might engage within the context of English and language arts

classrooms. Using a real-time, computer-based coding scheme, Martin Nystrand and Adam Gamoran recorded the amount of time spent in such activities as lectures, question-and-answer sessions, discussions, seatwork, small group work, student presentations, tests and quizzes, and so on. In general, such basic differences in how high- and low-track classrooms spend their time are relatively small and cannot explain the large differences in achievement growth over the course of the year.

As noted, however, there are differences in instruction that stem from teachers' and students' reactions to tracked learning environments, which help to explain track differences in achievement growth. Two broad instructional problems tend to plague instruction in low-track classrooms: reduced academic content and lack of engaging instruction.

Reduced Academic Content

Perhaps the clearest example of the reduced academic content of low-track learning environments comes from research on instruction in reading groups in early elementary school. In their studies of reading instruction, Rebecca Barr, Robert Dreeben, and Adam Gamoran measured the vocabulary development of first graders in different reading groups. They found a simple explanation for the reduced achievement growth in low-ability groups: the higher the mean aptitude of the reading group, the more words from the basal readers were likely to be covered, and consequently, more words were learned. This "words taught = words learned" conception of achievement growth is too simple to be applied directly to secondary school classrooms, but the basic insight—that tracking influences the nature of classroom instruction—can be elaborated to understand more nuanced differences in academic content.

"Ever notice how slooow this class is?" Page overheard one student remark to another in a low-track classroom. As Dreeben and colleagues found in first-grade classrooms, research on tracking often finds reduced content coverage. Sometimes, this is quite explicit; for example, many schools have a low-track algebra course designed to cover the same material as a high-track class, but over a 2-year period. It is important to note, however, that evidence shows that the majority of students in low-track math courses would benefit from more rigorous content. One reason for a slower pace and reduced

content coverage is that low-track classrooms spend somewhat less time actually engaged in instruction. The extra class time not spent in instruction is often spent disciplining students or engaged in "procedures and directions." Moreover, low-track classrooms often wander off-topic, reducing academic content further. For example, Nystrand and Gamoran found that high-track classrooms had about the same amount of discussion as low-track classrooms, but the content of low-track classrooms was often off-topic and nonacademic, whereas high-track classes engaged in focused discussions of academic material. Arthur Applebee and colleagues also found that classroom discourse was more closely tied to the curriculum in high-track English classrooms, and that the overall academic demand in classroom tasks was higher.

Less Engaging Instruction

Perhaps more important than the sheer amount of academic content in secondary school classrooms is whether this content is delivered in an engaging manner. Nystrand and Gamoran detail the rote nature of low-track teachers' approach to English instruction. Within the broad category of seatwork or Q&A, low-track classrooms engaged in activities such as filling in the blanks, answering true-false questions, and working on punctuation and grammar far more than high-track classrooms. One of the risks of such a highly structured approach, and the general preoccupation with order in low-track classrooms, is that students may find instruction less interesting and meaningful. Another risk of such instruction is that it is likely to be fragmented. In English classrooms, intertextuality—the process of alluding to another text during a textual analysis, or linking texts in order to better understand the literary elements of a text—is an important element of literature instruction. Page and Caughlan and Kelly found lower incidences of intertextuality in low-track classrooms, which contributed to a lack of coherence across lessons. Another important element of literature instruction that was missing in the low-track classroom was the linking of literature to students' lives, including their imagined future life as students, members of the workforce, and heads of households. Perhaps as a result of the highly structured but also fragmented instruction in low-track classrooms, researchers have observed that low-track students are frequently off-task. Using the National

Longitudinal Survey (base year 1988), William Carbonaro found that high-track students had higher levels of effort, in part because of beliefs that they were competent, but also because of more intellectually stimulating instruction.

It is important to note, however, that not all low-track teachers take such a teacher-centered approach to instruction. In fact, in Applebee et al.'s study, low-track classrooms were actually somewhat more likely than high-track classrooms to engage in imaginative writing and writing involving personal experience.

Future Directions

Tracking is a widespread educational process that, unfortunately, contributes significantly to educational inequality. Although some low-track classrooms are rich instructional environments, all too often students and teachers respond negatively to being in a low-track environment. The instruction that results has reduced academic content and proceeds in a manner that is less engaging to students. Despite the convincing evidence on the differences between high- and low-track classrooms, researchers disagree sharply about what is to be done about tracking. Some call for the wholesale detracking of schools, whereas others call for incremental reform of tracking systems to meet the intended goal of optimal instruction for all students.

There is not enough rigorous research on wholesale detracking to evaluate its effects on educational inequality. However, research addressing incremental reform of tracking has arisen in studies of sector differences in tracking systems, teacher tracking, and track assignment criteria. This research demonstrates that relatively small changes in the way tracking is enacted, such as encouraging all students to take more academic courses, can improve learning opportunities for low-track students. Moreover, the studies on instructional effects of tracking reviewed in this entry frequently contain exceptions to the rule, such as low-track classrooms where high achievement is the norm. It seems likely that the pernicious effects of tracking could be greatly reduced through a process of incremental change.

Sean Kelly

See also Motivation; Peer Influences; Self-Efficacy; Vocational Education

Further Readings

Abraham, J. (1989). Testing Hargreaves' and Lacey's differentiation-polarization theory in a setted comprehensive. *British Journal of Sociology*, *40*, 46–81.

Applebee, A. N., Langer, J. A., Nystrand, M., & Gamoran, A. (2003). Testing a framework for effective English instruction: A study of relationships between achievement gain and instructional approaches in middle and high school classrooms. *American Educational Research Journal*, *40*(3), 685–730.

Berends, M. (1995). Educational stratification and students' social bonding to school. *British Journal of Sociology of Education*, *16*, 327–351.

Carbonaro, W. (2005). Tracking, students' effort, and academic achievement. *Sociology of Education*, *78*, 27–49.

Caughlan, S., & Kelly, S. (2004). Bridging methodological gaps: Instructional and institutional effects of tracking in two English classes. *Research in the Teaching of English*, *39*, 20–62.

Gamoran, A. (1987). The stratification of high school learning opportunities. *Sociology of Education*, *60*, 135–155.

Kelly, S. (2004). Are teachers tracked? On what basis and with what consequences. *Social Psychology of Education*, *7*, 55–72.

Oakes, J. (1985). *Keeping track: How schools structure inequality.* New Haven, CT: Yale University Press.

Page, R. N. (1991). *Lower-track classrooms: A curricular and cultural perspective.* New York: Teachers College Press.

Schwartz, F. (1981). Supporting or subverting learning: Peer group patterns in four tracked schools. *Anthropology and Education Quarterly*, *12*, 99–119.

TRIARCHIC THEORY OF INTELLIGENCE

The triarchic theory of intelligence represents a way of understanding intelligence in broader terms than is the case for traditional theories.

Intelligence is defined in terms of the ability to achieve success in life in terms of one's personal standards, within one's sociocultural context. The field of intelligence has, at times, tended to put the cart before the horse, defining the construct conceptually on the basis of how it is operationalized rather than vice versa. This practice has resulted in tests that stress the academic aspect of intelligence, as one might expect, given the origins of modern intelligence testing in the work of Binet and Simon in designing an instrument that would distinguish children who would succeed from those who would fail in school. But the construct

of intelligence needs to serve a broader purpose, accounting for the bases of success in all of one's life.

One's ability to achieve success depends on one's capitalizing on one's strengths and correcting or compensating for one's weaknesses. Theories of intelligence typically specify some relatively fixed set of abilities, whether one general factor and a number of specific factors, seven multiple factors, eight multiple intelligences, or 150 separate intellectual abilities. Such a specification is useful in establishing a common set of skills to be tested. But people achieve success, even within a given occupation, in many different ways. For example, successful teachers and researchers achieve success through many different blendings of skills rather than through any single formula that works for all of them.

Balancing of abilities is achieved in order to adapt to, shape, and select environments. Definitions of *intelligence* traditionally have emphasized the role of adaptation to the environment. But intelligence involves not only modifying oneself to suit the environment (adaptation), but also modifying the environment to suit oneself (shaping), and sometimes, finding a new environment that is a better match to one's skills, values, or desires (selection).

Success is attained through a balance of analytical, creative, and practical abilities. Analytical abilities are the abilities primarily measured by traditional tests of abilities. But success in life requires one not only to analyze one's own ideas as well as the ideas of others, but also to generate ideas and to persuade other people of their value. This necessity occurs in the world of work, as when a subordinate tries to convince a superior of the value of his or her plan; in the world of personal relationships, as when a child attempts to convince a parent to do what he or she wants or when a spouse tries to convince the other spouse to do things his or her preferred way; and in the world of school, as when a student writes an essay arguing for a point of view.

Information-Processing Components

According to the proposed theory of human intelligence and its development, a common set of processes underlies all aspects of intelligence. These processes are hypothesized to be universal. For example, although the solutions to problems that are considered intelligent in one culture may be different from the solutions considered to be intelligent in another culture, the need to define problems and translate strategies to solve these problems exists in any culture.

Metacomponents, or executive processes, plan what to do, monitor things as they are being done, and evaluate things after they are done. Examples of metacomponents are recognizing the existence of a problem, defining the nature of the problem, deciding on a strategy for solving the problem, monitoring the solution of the problem, and evaluating the solution after the problem is solved.

Performance components execute the instructions of the metacomponents. For example, inference is used to decide how two stimuli are related, and application is used to apply what one has inferred. Other examples of performance components are comparison of stimuli, justification of a given response as adequate although not ideal, and actually making the response.

Knowledge-acquisition components are used to learn how to solve problems or simply to acquire declarative knowledge in the first place. Selective encoding is used to decide what information is relevant in the context of one's learning. Selective comparison is used to bring old information to bear on new problems. And selective combination is used to put together the selectively encoded and compared information into a single and sometimes insightful solution to a problem.

Although the same processes are used for all three aspects of intelligence universally, these processes are applied to different kinds of tasks and situations depending on whether a given problem requires analytical thinking, creative thinking, practical thinking, or a combination of these kinds of thinking. In particular, analytical thinking is invoked when components are applied to fairly familiar kinds of problems abstracted from everyday life. Creative thinking is invoked when the components are applied to relatively novel kinds of tasks or situations. Practical thinking is invoked when the components are applied to experience to adapt to, shape, and select environments.

Because the theory of intelligence comprises three subtheories—a componential subtheory dealing with the components of intelligence; an experiential subtheory dealing with the importance of coping with relative novelty and of automatization of information processing; and a contextual subtheory dealing with processes of adaptation, shaping, and selection—the theory has been referred to as *triarchic*.

Internal Validation

Internal validation seeks to show that a theory is internally consistent and that it is supported by data, independent of correlations with external measures.

Componential Analyses

Componential analyses involve studying the information-processing components underlying performance on cognitive tasks. These kinds of analyses have been used to study both analytical and creative thinking abilities.

Analytical Intelligence

Analytical intelligence is involved when the components of intelligence (which are specified by the componential subtheory of the triarchic theory) are applied to analyze, evaluate, judge, or compare and contrast. Typically, it is involved when components are applied to relatively familiar kinds of problems where the judgments to be made are of a fairly abstract nature.

In some early work, it was shown how analytical kinds of problems, such as analogies or syllogisms, can be analyzed componentially with response times or error rates decomposed to yield their underlying information-processing components. The goal of this research was to understand the information-processing origins of individual differences in (the analytical aspect of) human intelligence. With componential analysis, one could specify sources of individual differences underlying a factor score such as that for inductive reasoning. For example, response times on analogies and linear syllogisms were decomposed into their elementary performance components. The general strategy of such research is to (a) specify an information-processing model of task performance; (b) propose a parameterization of this model, so that each information-processing component is assigned a mathematical parameter corresponding to its latency (and another corresponding to its error rate); and (c) construct cognitive tasks administered in such a way that it is possible through mathematical modeling to isolate the parameters of the mathematical model.

Studies of reasoning need not use artificial formats. A more recent study explored predictions for everyday kinds of situations, such as when milk will spoil. The investigators looked at both predictions and postdictions (hypotheses about the past where information about the past is unknown) and found that postdictions took longer to make than did predictions.

Creative Intelligence. Intelligence tests contain a range of problems, some of them more novel than others. In some of the componential work, Sternberg and his colleagues have shown that when one goes beyond the range of unconventionality of the conventional tests of intelligence, one starts to tap sources of individual differences measured little or not at all by the tests. According to the theory of intelligence, (creative) intelligence is particularly well measured by problems assessing how well an individual can cope with relative novelty. Thus, it is important to include in a battery of tests problems that are relatively novel in nature. These problems can be either convergent or divergent in nature.

In work with convergent problems, Sternberg and his colleagues presented 80 individuals with novel kinds of reasoning problems that had a single best answer. For example, they might be told that some objects are green and others blue; but still other objects might be grue, meaning green until the year 2000 and blue thereafter, or bleen, meaning blue until the year 2000 and green thereafter. Or, they might be told of four kinds of people on the planet Kyron: blens, who are born young and die young; kwefs, who are born old and die old; balts, who are born young and die old; and prosses, who are born old and die young. Their task was to predict future states from past states, given incomplete information. In another set of studies, 60 people were given more conventional kinds of inductive reasoning problems, such as analogies, series completions, and classifications, but were told to solve them. But the problems had premises preceding them that were either conventional (dancers wear shoes) or novel (dancers eat shoes). The participants had to solve the problems as though the counterfactuals were true.

In these studies, Sternberg and his colleagues found that correlations with conventional kinds of tests depended on how novel or nonentrenched the conventional tests were. The more novel the items, the higher the correlations of the tests with scores on successively more novel conventional tests. Thus, the components isolated for relatively novel items would tend to correlate more highly with more unusual tests of fluid abilities than with tests of crystallized abilities. Sternberg and his colleagues also found that when response times on the relatively novel problems were componentially analyzed, some components measured the creative aspect of intelligence better than others. For example, in the "grue-bleen" task

mentioned above, the information-processing component requiring people to switch from conventional green-blue thinking to grue-bleen thinking and then back to green-blue thinking again was a particularly good measure of the ability to cope with novelty.

Componential analyses provide one means of internal validation of the triarchic theory. But their emphasis is on testing specific models of task performance for particular components of information processing. Is it possible internally to validate the triarchic theory as a whole?

Factor Analyses

Internal Validity

Several studies have looked at and been supportive of the factorial structure of the triarchic theory. In the most recent study by Sternberg and the Rainbow Project Collaborators, factors were found corresponding to creative, practical, and paper-and-pencil analytical measures of intelligence.

External Validation

External validation seeks to show that a theory has convergent and discriminant validity—that measures based on it predict what they are supposed to predict and do not predict what they are not supposed to predict.

Correlational Studies

Tests of successful intelligence predict success. In one study, tests of creative and practical intelligence doubled prediction of college grades obtained when the SAT was used alone. The tests are also useful when measuring abilities separately.

Analytical Intelligence

In the componential analysis work described above, correlations were computed between component scores of individuals and scores on tests of different kinds of psychometric abilities. The correlations obtained for all the components showed convergent-discriminant validation: They tended to be significant with psychometric tests of reasoning but not with psychometric tests of perceptual speed. Moreover, significant correlations with vocabulary tended to be obtained only for encoding of verbal stimuli.

Creative Intelligence

In work with divergent reasoning problems having no one best answer, the investigators asked 63 people to create various kinds of products where an infinite variety of responses was possible. Individuals were asked to create products in the realms of writing, art, advertising, and science. In writing, they would be asked to write very short stories for which the investigators would give them a choice of titles, such as "Beyond the Edge" or "The Octopus's Sneakers." In art, the participants were asked to produce art compositions with titles such as "The Beginning of Time" or "Earth From an Insect's Point of View." In advertising, they were asked to produce advertisements for products such as a brand of bow tie or a brand of doorknob. In science, they were asked to solve problems such as one asking them how people might detect extraterrestrial aliens among us who are seeking to escape detection. Participants created two products in each domain.

Sternberg and Lubart found, first, that creativity comprises the components proposed by their investment model of creativity: intelligence, knowledge, thinking styles, personality, and motivation. Second, they found that creativity is relatively, although not wholly, domain-specific. Correlations of ratings of the creative quality of the products across domains were lower than correlations of ratings and generally were at about the 0.4 level. Thus, there was some degree of relation across domains at the same time that there was plenty of room for someone to be strong in one or more domains but not in others. Third, Sternberg and Lubart found a range of correlations of measures of creative performance with conventional tests of abilities. As was the case for the correlations obtained with convergent problems, correlations were higher to the extent that problems on the conventional tests were nonentrenched. For example, correlations were higher with fluid than with crystallized ability tests, and the higher the correlations, the more novel the fluid test was. These results show that tests of creative intelligence have some overlap with conventional tests (e.g., in requiring verbal skills or the ability to analyze one's own ideas) but also tap skills beyond those measured even by relatively novel kinds of items on the conventional tests of intelligence.

The work on creativity revealed a number of sources of individual and developmental differences.

To what extent was the thinking of the individual novel or nonentrenched?

What was the quality of the individual's thinking?

To what extent did the thinking of the individual meet the demands of the task?

Tests of creative intelligence go beyond tests of analytical intelligence in measuring performance on tasks that require individuals to deal with relatively novel situations. At the same time, they probably measure creativity that is, for the most part, within existing paradigms. But how about situations that are relatively familiar, but in a practical rather than an academic domain? Can one measure intelligence in the practical domain, and if so, what is its relation to intelligence in more academic kinds of domains?

Practical Intelligence

Practical intelligence involves individuals applying their abilities to the kinds of problems that confront them in daily life, such as on the job or in the home. Practical intelligence involves applying the components of intelligence to experience so as to (a) adapt to, (b) shape, and (c) select environments. Adaptation is involved when one changes oneself to suit the environment. Shaping is involved when one changes the environment to suit oneself. And selection is involved when one decides to seek out another environment that is a better match to one's needs, abilities, and desires. People differ in their balance of adaptation, shaping, and selection, and in the competence with which they balance among the three possible courses of action.

Much of the work of Sternberg and his colleagues on practical intelligence has centered on the concept of tacit knowledge. They have defined this construct as what one needs to know in order to work effectively in an environment that one is not explicitly taught and that often is not even verbalized.

Sternberg and colleagues typically have measured tacit knowledge using work-related problems that present problems one might encounter on the job. They have measured tacit knowledge for both children and adults, and among adults, for people in more than two dozen occupations, such as management, sales, academia, teaching, school administration, secretarial work, and the military.

In the tacit-knowledge studies, Sternberg and his colleagues have found, first, that practical intelligence as embodied in tacit knowledge increases with experience, but it is profiting from experience, rather than experience per se, that results in increases in scores. Some people can be in a job for years and still have acquired relatively little tacit knowledge. Second, they also have found that subscores on tests of tacit knowledge—such as for managing oneself, managing others, and managing tasks—correlate significantly with each other. Third, scores on various tests of tacit knowledge, such as for academics and managers, are also correlated fairly substantially (at about the 0.5 level) with each other. Thus, fourth, tests of tacit knowledge may yield a general factor across these tests. However, fifth, scores on tacit-knowledge tests do not correlate with scores on conventional tests of intelligence, whether the measures used are single-score measures of multiple-ability batteries. Thus, any general factor from the tacit-knowledge tests is not the same as any general factor from tests of academic abilities (suggesting that neither kind of g factor is truly general, but rather, general only across a limited range of measuring instruments). Sixth, despite the lack of correlation of practical intellectual with conventional measures, the scores on tacit-knowledge tests predict performance on the job as well as or better than do conventional psychometric intelligence tests. Seventh, in one study done at the Center for Creative Leadership, they further found that scores on tests of tacit knowledge for management were the best single predictor of performance on a managerial simulation.

In work on military leadership, Hedlund and colleagues found, eighth, that scores of 562 participants on tests of tacit knowledge for military leadership predicted ratings of leadership effectiveness, whereas scores on a conventional test of intelligence and on a tacit-knowledge test for managers did not significantly predict the ratings of effectiveness.

Culture also matters in successful intelligence. In a study in Usenge, Kenya, near the town of Kisumu, Sternberg and his colleagues were interested in school-age children's ability to adapt to their indigenous environment. They devised a test of practical intelligence for adaptation to the environment. The test of practical intelligence measured children's informal tacit knowledge for natural herbal medicines that the villagers believe can be used to fight various types of infections. The researchers measured the Kenyan children's ability to identify the medicines, where they come from, what they are used for, and how they are dosed. Based on work the

researchers had done elsewhere, they expected that scores on this test would not correlate with scores on conventional tests of intelligence. In order to test this hypothesis, they also administered to the 85 children the Raven Coloured Progressive Matrices Test, which is a measure of fluid- or abstract-reasoning-based abilities, as well as the Mill Hill Vocabulary Scale, which is a measure of crystallized- or formal-knowledge-based abilities. In addition, they gave the children a comparable test of vocabulary in their own Dholuo language. The Dholuo language is spoken in the home, English in the schools.

The researchers did indeed find no correlation between the test of indigenous tacit knowledge and scores on the fluid-ability tests. But to their surprise, they found statistically significant correlations of the tacit-knowledge tests with the tests of crystallized abilities. The correlations, however, were *negative*. In other words, the higher the children scored on the test of tacit knowledge, the lower they scored, on average, on the tests of crystallized abilities.

Instructional Studies

Researchers explored the question of whether conventional education in school systematically discriminates against children with creative and practical strengths. Motivating this work was the belief that the systems in most schools strongly tend to favor children with strengths in memory and analytical abilities. Also motivating the work was an opportunity externally to validate the triarchic theory of intelligence.

The investigators used the Sternberg Triarchic Abilities Test, as described above, in some of their instructional work. The test was administered to 326 children around the United States and in some other countries who were identified by their schools as gifted by any standard whatsoever. Children were selected for a summer program in (college-level) psychology if they fell into one of five ability groupings: high analytical, high creative, high practical, high balanced (high in all three abilities), or low balanced (low in all three abilities). Students who came to Yale University were then divided into four instructional groups. Students in all four instructional groups used the same introductory psychology textbook and listened to the same psychology lectures. What differed among them was the type of afternoon discussion section to which they were assigned. They were assigned

to an instructional condition that emphasized either memory, analytical, creative, or practical instruction. For example, in the memory condition, they might be asked to describe the main tenets of a major theory of depression. In the analytical condition, they might be asked to compare and contrast two theories of depression. In the creative condition, they might be asked to formulate their own theory of depression. In the practical condition, they might be asked how they could use what they had learned about depression to help a friend who was depressed.

Students in all four instructional conditions were evaluated in terms of their performance on homework, a midterm exam, a final exam, and an independent project. Each type of work was evaluated for memory, analytical, creative, and practical quality. Thus, all students were evaluated in exactly the same way.

All three ability tests—analytical, creative, and practical—significantly predicted course performance. When multiple regression analysis was used, at least two of these ability measures contributed significantly to the prediction of each of the measures of achievement. Perhaps as a reflection of the difficulty of deemphasizing the analytical way of teaching, one of the significant predictors was always the analytical score. Most importantly, there was an aptitude-treatment interaction whereby students who were placed in instructional conditions that better matched their pattern of abilities outperformed students who were mismatched. In other words, when students are taught in a way that fits how they think, they do better in school. Children with creative and practical abilities, who are almost never taught or assessed in a way that matches their pattern of abilities, may be at a disadvantage in course after course, year after year.

A follow-up study by Sternberg, Torff, and Grigorenko examined learning of social studies and science by third graders and eighth graders. The 225 third graders were students in a very low-income neighborhood in Raleigh, North Carolina. The 142 eighth graders were students who were largely middle class to upper middle class studying in Baltimore, Maryland, and Fresno, California. In this study, students were assigned to one of three instructional conditions. In the first condition, they were taught the course that basically they would have learned had there been no intervention. The emphasis in the course was on memory. In a second condition, students were taught in a way that emphasized critical (analytical) thinking. In the

third condition, they were taught in a way that emphasized analytical, creative, and practical thinking. All students' performance was assessed for memory learning (through multiple-choice assessments) as well as for analytical, creative, and practical learning (through performance assessments).

As expected, students in the successful-intelligence (analytical, creative, practical) condition outperformed the other students in terms of the performance assessments. One could argue that this result merely reflected the way they were taught. Nevertheless, the result suggested that teaching for these kinds of thinking succeeded. More important, however, was the result that children in the successful-intelligence condition outperformed the other children even on the multiple-choice memory tests. In other words, to the extent that one's goal is just to maximize children's memory for information, teaching for successful intelligence is still superior. It enables children to capitalize on their strengths and to correct or compensate for their weaknesses, and it allows children to encode material in a variety of interesting ways.

Grigorenko and her colleagues extended these results to reading curricula at the middle school and the high school level. In a study of 871 middle school students and 432 high school students, researchers taught reading either triarchically or through the regular curriculum. At the middle school level, reading was taught explicitly. At the high school level, reading was infused into instruction in mathematics, physical sciences, social sciences, English, history, foreign languages, and the arts. In all settings, students who were taught triarchially substantially outperformed students who were taught in standard ways.

Thus, the results of three sets of studies suggest that the triarchic theory of intelligence is valid as a whole. Moreover, the results suggest that the theory can make a difference not only in laboratory tests, but in school classrooms and even the everyday lives of adults as well.

Robert J. Sternberg

See also Intelligence and Intellectual Development; Intelligence Quotient (IQ); Intelligence Tests; Learning

Further Readings

Binet, A., & Simon, T. (1916). *The development of intelligence in children.* Baltimore, MD: Williams & Wilkins. (Originally published in 1905)

Cattell, R. B., & Cattell, H. E. P. (1973). *Measuring intelligence with the Culture Fair Tests.* Champaign, IL: Institute for Personality and Ability Testing.

Gardner, H. (1999). *Intelligence reframed: Multiple intelligences for the 21st century.* New York: Basic Books.

Grigorenko, E. L., Jarvin, L., & Sternberg, R. J. (2002). School-based tests of the triarchic theory of human intelligence: Three settings, three samples, three syllabi. *Contemporary Educational Psychology, 27,* 167–208.

Hedlund, J., Forsythe, G. B., Horvath, J. A., Williams, W. M., Snook, S., & Sternberg, R. J. (2003). Identifying and assessing tacit knowledge: Understanding the practical intelligence of military leaders. *Leadership Quarterly, 14,* 117–140.

Spearman, C. (1904). "General intelligence," objectively determined and measured. *American Journal of Psychology, 15*(2), 201–293.

Sternberg, R. J. (1997). *Successful intelligence.* New York: Plume.

Sternberg, R. J. (2004). Culture and intelligence. *American Psychologist, 59*(5), 325–338.

Sternberg, R. J., & Detterman, D. K. (Eds.). (1986). *What is intelligence? Contemporary viewpoints on its nature and definition.* Norwood, NJ: Ablex.

Sternberg, R. J., & Gastel, J. (1989). Coping with novelty in human intelligence: An empirical investigation. *Intelligence, 13,* 187–197.

Sternberg, R. J., Grigorenko, E. L., Ferrari, M., & Clinkenbeard, P. (1999). The triarchic model applied to gifted identification, instruction, and assessment. In N. Colangelo & S. G. Assouline (Eds.), *Talent development III: Proceedings from the 1995 Henry B. and Jocelyn Wallace National Research Symposium on Talent Development* (pp. 71–80). Scottsdale, AZ: Gifted Psychology Press.

Sternberg, R. J., & The Rainbow Collaborators. (2006). Augmenting the SAT through assessments of analytic, practical, and creative skills. In W. J. Camara & E. W. Kimmel (Eds.), *Choosing students: Higher education tools for the 21st century.* Mahwah, NJ: Lawrence Erlbaum.

Sternberg, R. J., Torff, B., & Grigorenko, E. L. (1998). Teaching triarchically improves school achievement. *Journal of Educational Psychology, 90*(3), 374–384.

Tetewsky, S. J., & Sternberg, R. J. (1986). Conceptual and lexical determinants of nonentrenched thinking. *Journal of Memory and Language, 25,* 202–225.

V

It is not more vacation we need—it is more vocation.

—Eleanor Roosevelt

VALIDITY

According to the *Standards for Educational and Psychological Testing*, published in 1999 by the American Educational Research Association, American Psychological Association, and the National Council on Measurement in Education, *validity* is defined as the extent to which empirical evidence and theory lend support to the interpretation and inferences made about test scores for particular uses. Validity is considered the most fundamental component of developing and evaluating educational and psychological tests. This is understandable. Validity involves the accumulation of evidence and theoretical justification for claims made about the traits, abilities, or attributes (e.g., intelligence, achievement, self-esteem) individuals might possess as measured by educational and psychological tests. In the following three sections, the process of validating tests will be reviewed. In the first section, the general approach to argument-based validation will be described, including a description of how test-related aspects can undermine this argument. In the second section, the sources of validity evidence—test content, response processes, internal structure, relations to other variables, and consequences—normally collected to bolster validity claims will be outlined. In the third section, the methods for integrating sources of evidence to build a strong validity argument are explained. This entry will reflect much of the information contained in the *Standards for Educational and Psychological Testing*, which, in turn reflects the scientific debates, contributions, and consensus of leading educational and psychological researchers and practitioners in the field.

Argument-Based Approach to Validation

Educational and psychological tests are pervasive in society, offering measurements of many distinct attributes, such as academic aptitude, personality characteristics, and career interests. The responsible use of tests can lead to increased information about individuals so as to help them lead more productive, happier lives. However, when tests are used to make unwarranted claims about individuals, it hurts not only the persons writing the test but also the general public, who may question the development and use of tests for the good of society.

Establishing the validity of inferences made about individuals (henceforth referred to as *examinees*) based on their test scores is fundamental to the appropriate development and use of tests.

The first step toward establishing the validity of test-score interpretations for particular uses is a written statement or series of written statements about how the test scores are going to be interpreted. The argument-based approach to validation requires that

these statements be clear and explicit and also include a rationale for the proposed interpretation and use. For example, imagine a developer of a science achievement test who intends to use the test to select high school students for a special university program in physics. The developer would begin the validation process by writing an explicit statement about how the test scores on the achievement test are going to be interpreted; for example, high scores will be interpreted to indicate competence in the following knowledge and skills: designing experiments, testing hypotheses, and deductive reasoning. The statement would also indicate the proposed uses of test-scores; for example, examinees scoring above a certain cut-off will be selected for inclusion in the university physics program. A rationale for viewing the test items, and hence test scores, as appropriate measures of the construct of interest (science achievement) and why the scores should lead to a particular use must also be outlined. For example, the rationale must provide reasons for accepting the items as appropriate measures of the knowledge and skills in science achievement and why these knowledge and skills (and not others) are the appropriate ones to use in the selection of examinees.

Some categories of evidence are better than others to help generate and develop a strong argument for the proposed interpretation and use of test scores. For example, in determining whether the science achievement test mentioned in the previous example is appropriate for selecting examinees to participate in the university program in physics, the test developer needs to consider whether (a) the knowledge and skills measured by the test are those that will be necessary to succeed and profit from the university physics course; (b) the test does indeed measure those knowledge and skills deemed relevant, and the test does discriminate between those students who will and will not profit from taking part in the university physics course; (c) the test scores can be generalized to indicate how students will perform on similar sets of items; (d) the test scores are not influenced by secondary variables such as test-wiseness, which can inflate performance on the test but for reasons that are unrelated to the attributes or construct of interest; (e) student success in the program in physics can be measured so as to determine whether the test was useful in selecting examinees for the program; and (f) it is true that examinees who perform very well on the science achievement test will be more successful in the university course than students who do not perform well on the science test.

In addition to considering these categories of evidence in developing a strong argument for the proposed interpretation and use of test scores, it is also useful to consider rival hypotheses, namely, what other interpretations and inferences could be drawn from the test scores. In trying to imagine what these other interpretations might be, the developer is encouraged to consider the perspectives of different groups, such as parents, administrators, and counselors, who may have experience with similar types of tests, and the potential consequences associated with the proposed interpretation and use of test scores for examinees and other stakeholders.

When consulting with stakeholders, it may become apparent that the test is measuring a construct that is not fully specified, meaning that it fails to include important aspects of, for example, science achievement. Conversely, the test may include nuisance variables that may lead to unwanted variance in test scores. In particular, when considering the perspectives of others, developers are encouraged to consider two potential sources of bias that can undermine the argument for validity: *construct underrepresentation* and *construct-irrelevant variance*. Both sources of bias are problematic.

In the first case, *construct underrepresentation* means that a test does not measure important attributes associated with the construct and therefore will lead to problematic test-score interpretations. For example, the science achievement test may be designed to measure deductive reasoning but may not contain any items measuring inductive reasoning. If test scores will be interpreted to indicate how well examinees have mastered the knowledge and skills within the science domain, there is a problem with not including a measure of inductive reasoning. Inductive reasoning is considered an important skill in the generation of hypotheses. Thus, using the test scores to make inferences about science achievement will be problematic because the construct of the science achievement test is underrepresented; that is, inductive reasoning has been left out.

In the second case, *construct-irrelevant variance* means that a test is measuring nuisance attributes that will not be considered when interpreting test scores and may, in fact, disrupt the measurement of the real attributes of interest. For example, imagine the science achievement test includes many pictorial illustrations. The illustrations are included in the test for aesthetic purposes but do provide useful information

to examinees as they answer the test items. Although some examinees may not be distracted by the illustrations, others could be distracted, thinking that the illustrations are relevant to responding to the items. As a result, some examinees might spend time studying the illustrations only to find out that they have run out of time to complete the test. If the illustrations result in these examinees performing poorly in spite of their achievement of scientific knowledge and skills, there is a problem with using these test scores for the purpose intended.

The process of validation is considered the responsibility of both the developer of the test and the user. Test developers must bring evidence to bear to the interpretation of test scores for a particular use. In addition to pointing to the relevant literature, collecting empirical evidence, and showing the soundness of conclusions with respect to the interpretation of test scores, test developers also must show that even more mundane aspects of the test—for example, test and item format, the conditions of the test administration, the difficulty of language used—do not bias the interpretation of test scores. Likewise, the test user also has responsibilities. The test user who selects to use the test in a particular situation must show that the test is being used in a manner that is consistent with the purpose proposed by the test developer. If the test user intends to use the test in a new way, that is, a way that is dissimilar to the original purpose proposed by the test developer, the user must bring evidence to bear for the new purpose. In other words, the user becomes responsible for demonstrating that the new purpose is empirically and logically consistent with the construct measured by the test. In the next section, the sources of evidence that are brought to bear to the interpretation of test scores are considered.

Sources of Validity Evidence

There are five sources of validity evidence. These sources include evidence based on test content material, examinees' response processes, internal test structure, relationships of the test to other variables, and consequences of testing.

Test Content

Test content refers to the thematic features of the test, such as the material included in the test, the types of words used to ask test questions, the format of the

questions (e.g., open-ended or multiple-choice), the directions communicated to examinees as to how to respond to questions, and even the procedures used to administer the test and scoring. Test content can serve as a source of evidence in the argument for validity. Toward this end, the test developer must demonstrate that the content included in the test is related to the construct being measured. For example, a developer designing a test of reading comprehension to measure examinees' mastery of the reading curriculum should demonstrate that the content of the test (e.g., reading passages, questions posed), the format of the items (e.g., multiple-choice, open-ended), the directions to examinees, and other content-related aspects of the test are aligned with measuring mastery of the reading curriculum in the program of study.

Often developers will seek the judgment of experts to furnish evidence about test content. For example, a panel of individuals with expertise in language arts, the psychology of reading behavior, and test design may be convened to decide on the adequacy of the specifications used to select and design the items for the reading comprehension test. The specifications may include a two-way matrix showing the content themes to be covered in the test along with reading comprehension skills. The panel of experts might judge the adequacy of the specifications in relation to the program of study, including the match between the themes and skills in the specifications with those in the program, their frequency, and the suitability of the test questions in terms of measuring the themes and skills of interest.

Response Processes

Evidence based on response processes requires gathering information about how examinees think about the content of test items and how they use this content to generate responses. Evidence based on response processes is necessary when test scores will be interpreted to indicate examinees' proficiency in the type of thinking that led to a response, such as cognitive skill in manipulating information, combining ideas, generating strategies, or reasoning from premises to conclusion. For example, a test designed to measure critical reasoning and thinking must include evidence to show that examinees who score highly on the test do indeed weigh alternate perspectives, identify conflicts of interest, question unsupported conclusions, and recognize bias. In other words, if test scores will be interpreted to

indicate proficiency in critical reasoning and thinking, evidence must be brought forward to show that examinees who score highly on the test do reason and think critically and do not simply generate correct answers by following a rote algorithm.

Evidence on response processes can be obtained by conducting empirical studies with representative samples of examinees, such as interviewing examinees and asking them questions about how they arrived at their answers, tracking their eye movements as they read and respond to test items, and measuring response times. The specific type of empirical evidence collected will be determined by the construct measured by the test and the interpretation of test scores. For example, the developer of a diagnostic test of reading ability might collect eye-tracking information from examinees if he or she intends to interpret test scores as indicating the reading difficulty examinees experience with certain parts of the text. Evidence of response processes can also be used to support test-score interpretations across groups of examinees. For example, in order for test scores to be interpreted similarly for both boys and girls, there needs to be evidence that boys and girls think similarly about the test questions and that high or low scores for the groups reflect equivalent response processes.

Internal Structure

Evidence based on the internal structure of a test usually involves generating quantitative information about the test, that is, evidence about the relationships among test items. For example, consider a test designed to measure verbal fluency. If the construct of the test is assumed to be integrated and not divisible to component parts, and the test items subsequently developed to measure verbal fluency are homogenous and designed to differ only in difficulty, one would expect the test to have simple structure and for there to be a single statistical dimension or latent trait to account for variability in examinee test performance. Test scores could then be interpreted as indicating the amount of verbal fluency along a single, unitary scale. In an exploratory factor analytic study of test item responses, one would expect to identify only a single factor accounting for test performance. If this single factor were not found and instead the test of verbal fluency yielded two factors, the failure to identify only a single factor would undermine the original, proposed interpretation of test scores. Verbal fluency

items measuring *oral fluency* specifically might load on one factor and other items measuring *written fluency* specifically might load on a second factor, suggesting that it may not be useful to interpret verbal fluency along a single scale but to interpret test scores using at least two scales.

Other considerations related to the internal structure of a test, especially a test expected to have complex structure, include the number of dimensions expected to underlie the test and the correlations among the identified dimensions. Evidence can also be gathered about how test items function in different populations of interest. For example, differential item functioning occurs when different groups of examinees, such as boys and girls, with identical overall ability on an appropriate criterion have different probabilities of answering a test item correctly. This can happen when the content of test items is more familiar to boys or girls, thus offering one of the groups an unfair advantage in answering the item correctly irrespective of ability. Differential item functioning is not always attributable to bias, but test items should be examined closely for their potential to produce differential item functioning.

Relations to Other Variables

When collecting evidence based on the relation of test scores to other variables, the developer often collects evidence about how the scores of one test correlate with another measure that is believed to function as a criterion. The criterion measure is considered of primary interest. For example, the Scholastic Assessment Test (SAT), administered by the College Board and scored by the Educational Testing Service, was designed as a predictor of college success, with first-year college grade point average viewed as the relevant criterion measure of college success. Because the criterion variable—college success—is what the test aims to predict, this means that how well SAT test scores correlate with first-year college grades becomes an important source of validity evidence to show that the test is working as it should. If SAT test scores were not correlated with first-year college grades, the absence of this relationship would undermine the aim of the SAT for predicting college success.

Test scores are often correlated with other variables believed to measure similar or dissimilar constructs. When a test is correlated with another test, which is believed to measure the same construct, the type of

evidence generated by the correlation is called *convergent evidence.* For example, a new intelligence test designed to be administered in less time and scored with greater efficiency is correlated with test scores on the well-established Stanford–Binet test to show that the new test measures intelligence. If the scores on the new intelligence test correlate highly with the scores on the Stanford–Binet, the developers of the new test could use this evidence to show that the new test is indeed measuring intelligence. When test scores are correlated with measures of dissimilar constructs, then much lower correlations are expected. This type of evidence is called *discriminant evidence* and is designed to show that the construct measured by one test is distinct from other constructs measured by other tests. For example, one might not expect a test of intelligence to be correlated with measures of physical activity or personality. Convergent and discriminant evidence can also be obtained from experimental studies.

Consequences

The consequences associated with test use can be relevant to validation. This is especially true when the consequences are traced to sources of bias in the test, leading to unintended and negative consequences for members of particular groups. For example, consider a job-placement test designed to select new candidates for managerial positions. Imagine that the test requires potential candidates to solve advanced algebra questions even though the managerial positions the test is designed to help fill do not require individuals to perform advanced mathematics and only require basic computation skills. If there is no obvious rationale for including the advanced algebra questions on the test, these questions add construct-irrelevant variance to test scores and may disadvantage certain groups from being selected to be managers. In this case, the test may be biased and may unfairly jeopardize a group of examinees. The issue of whether consequences are relevant to the process of validation has been controversial, leading many researchers and practitioners to focus on issues related to disentangling validity from fairness in testing.

Recent Advances to the Argument-Based Approach to Validation

The modern view of validity in educational and psychological testing has a rich and complex history, with many of its foundations found in the early writings of Lee Cronbach, Jane Loevinger, and others. However, Michael Kane advances the most current treatment of the argument-based validation framework. According to Kane, the argument-based validation framework begins with the *interpretative argument.* The interpretative argument includes an explicit set of statements about the proposed interpretation and uses for a set of test scores and provides the unequivocal chain of reasoning supporting the interpretation and uses proposed. In this chain of reasoning, the test developer illustrates the relevant empirical observations and other sources of evidence that are being used to generate, support, and defend the proposed interpretations and uses.

The interpretive argument lays out the validity evidence in favor of the proposed interpretation and uses of test scores. The interpretive argument needs to be critically evaluated for whether it provides compelling reason to endorse the interpretation and uses proposed. This is where the *validity argument* is initiated. The validity argument is initiated at the end of the validation process as a way to critically evaluate the contents and chain of reasoning offered in the interpretive argument. In evaluating the interpretive argument, stakeholders must determine and evaluate the coherence of the argument, that is, the rigor of the empirical evidence presented, the reasonableness of the inferences and supporting assumptions made about the evidence, and the overall strength of the conclusions generated based on the evidence and associated inferences. In general, empirical studies of the most controversial or questionable suppositions are likely to be most informative. Moreover, the inferences in the chain of reasoning will often take the form of *if-then* rules, such as *if the examinee has not mastered knowledge and skills X and Y, then the observed score on the test should be this value.* The supporting assumptions included in the chain of reasoning should provide support for the inferences made within the interpretative argument. For example, a supporting assumption might be that an examinee's sample of performance on the test items is representative of his or her performance on the universe of items. A summary of the steps involved in the argument-based approach is provided here:

1. Identifying a written statement of the proposed interpretation of test scores and test use

2. Designing a plan to generate evidence of the proposed interpretation and test use, including framework for

empirical evidence, inferences, and assumptions of the interpretive argument

3. Developing the test instrument, including efforts to identify and control sources of bias

4. Evaluating the inferences and assumptions of the interpretative argument during test development, in which modifications to the testing instrument are made

5. Critically appraising the interpretive argument

6. Searching for hidden assumptions in the interpretative argument and determining the strength of evidence for questionable suppositions

7. Determining persuasiveness of argument for different groups of stakeholders

One of the immediate virtues of the argument-based validation framework is the transparency it provides to stakeholders about the nature of the evidence used to interpret test scores and decide upon test use. Through the argument-based validation framework, test developers can explain their reasoning for the meaning of test scores and actions for test use. Stakeholders may scrutinize the premises and conclusion of the interpretative argument and determine whether it reflects cohesion of thought and leads to strong conclusions.

In closing, validity is defined as the extent to which empirical evidence and theory lend support to the interpretation and inferences made about test scores for particular uses. In establishing the validity of test-score interpretations and uses, an argument-based framework is invoked. This framework involves developing an interpretative argument, which is then critically evaluated (the validity argument) for coherence and strength. In the interpretative argument, sources of empirical evidence, along with associated inferences and supporting assumptions, are presented. Sources of empirical evidence can include test content, examinees' response processes, internal test structure, test item relations to other variables, and consequences. Moreover, the chain of reasoning used to combine empirical evidence and supporting assumptions is explicitly laid out to generate a transparent rationale for adopting the proposed test score interpretations for particular uses. The appraisal of the interpretative argument is scrutinized for its coherence, reasonableness, and strength.

Jacqueline P. Leighton

See also Assessment; Reliability; Testing

Further Readings

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing.* Washington, DC: Author.

Borsboom, D., Mellenbergh, G. J., & Van Heerden, J. (2004). The concept of validity. *Psychological Review, 111,* 1061–1071.

Cronbach, L. J. (1971). Test validation. In R. L. Thorndike (Ed.), *Educational measurement* (2nd ed., pp. 443–507). Washington, DC: American Council on Education.

Cronbach, L. J. (1988). Five perspectives on validation argument. In H. Wainer & H. Braun (Eds.), *Test validity* (pp. 3–17). Hillsdale, NJ: Lawrence Erlbaum.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52,* 281–302.

Cureton, E. E. (1951). Validity. In E. F. Lindquist (Ed.), *Educational measurement* (pp. 621–694). Washington, DC: American Council on Education.

Kane, M. T. (2006). Validation. In R. Brennan (Ed.), *Educational measurement* (4th ed., pp. 17–64). Washington, DC: American Council on Education.

Loevinger, J. (1957). Objective tests as instruments of psychological theory. *Psychological Reports, 3*(Suppl 9), 635–694.

Messick, S. (1989) Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 1–103). New York: American Council on Education, Macmillan.

Suppe, P. (1977). *The structure of scientific theories.* Urbana: University of Illinois Press.

VICARIOUS REINFORCEMENT

Vicarious reinforcement occurs when (a) an individual observes another person (a model) behave in a certain way and experience a consequence perceived as desirable by the observer, and (b) as a result, the observer behaves as the model did. For example, suppose a shy child at school observes another student being praised by the teacher for speaking up in class. The observed student is the reinforced model. If the shy child would like to be praised by the teacher and therefore personally speaks up in class in the future, vicarious reinforcement has occurred. Educators commonly use vicarious reinforcement to shape the behavior of students.

Psychologist Albert Bandura popularized the concept of vicarious reinforcement through his theory of human behavior, *social cognitive theory* (sometimes called *social learning theory*). Social cognitive theory proposes that human behavior is the outcome of

interactions of cognitions, behavior, and environment. An important environmental influence on an individual's behavior is the model set by others. Social cognitive theory suggests that a great deal of learning occurs through observation of a model's behavior and the consequences of the behavior. Bandura labeled this phenomenon *observational learning.*

Researchers of social learning have demonstrated the effects of vicarious reinforcement through experimental manipulations. For example, studies involving children receiving social reinforcement (praise) for their performances on experimental tasks have shown that other children who observe the social reinforcement try harder at the tasks.

Vicarious reinforcement can also occur symbolically. That is, rather than actually observing the reinforced behavior occurring, individuals can experience vicarious reinforcement by merely acquiring knowledge about another individual who has performed a behavior and benefited. Examples of this include (a) hearing a personal anecdote from an individual who has benefited from engaging in a behavior and (b) observing an individual receiving an award for an outstanding performance. Some schools systematically apply symbolic vicarious reinforcement by giving individual students awards at student assemblies or commending the students in the school newsletter.

Studies of symbolic vicarious reinforcement have shown that it can affect behavior. For instance, one study showed that vicarious reinforcement, delivered symbolically via personal anecdotes from individuals on a video, increased observers' use of a recommended method of coping with stressors.

There are three conditions that enhance the likelihood of vicariously rewarded behavior being imitated by an observer: (1) the model is liked, (2) the model is similar to the observer (e.g., the same sex), and (3) the model is of higher status than the observer. Applying these principles to increase students' efforts to work hard in class, teachers might praise a well-liked student who has just worked hard. It might help if the teachers praised a boy as well as a girl and a good student as well as a struggling student so that most students would see the praised student as similar in important ways to themselves. Finally, the teachers might talk about how hard they worked in class when they were students and thereby use themselves as high-status models.

Vicarious reinforcement works, at least in part, through expectancies similar to those involved in *operant conditioning* and *rule-governed behavior.* Operant conditioning involves the increase or reduction of a behavior as a result of its consequences. For example, if a student receives a compliment for working hard on a school assignment, the student will tend to associate hard work with praise and feeling good and will be more likely to work hard on assignments in the future. Rule-governed behavior occurs when someone simply tells a person that the person will receive a reward for a specific behavior, and as a result, the person shows the behavior. For example, a teacher might tell the class that everyone who submits homework on time will receive 15 minutes to play an exciting computer game. If students complete the homework because of their expectation of obtaining the computer time, they are engaging in rule-governed behavior. The expectancies involved in operant conditioning and rule-governed behavior play a similar role in vicarious reinforcement. When individuals learn through observation or verbal description about specific behavior leading to positive consequences, they tend to expect positive consequences if they engage in the behavior.

An important question about vicarious reinforcement is how it compares with direct reinforcement. In answering this, Bandura distinguished between *learning* and *motivational* effects. Compared with personally experienced outcomes, vicarious outcomes tend to produce stronger learning effects, perhaps because observers' attentional resources are not divided between performing behaviors and learning their outcomes. However, direct reinforcement is more effective than vicarious reinforcement in creating motivational effects, such as an actual increase in the rate of a behavior. Hence, an individual might learn the most about how to start a conversation with others by observing someone do that successfully, but actually starting a conversation successfully might be more likely to increase the rate of that behavior in the future.

On the flip side of vicarious reinforcement is *vicarious punishment.* Like vicarious reinforcement, vicarious punishment involves learning about the consequences of behaviors through observation or anecdote. However, vicarious punishment serves to deter observers from imitating a modeled behavior because the behavior has resulted in a negative outcome for the model. Teachers sometimes create vicarious punishment when they reprimand a misbehaving student in front of other students or when they tell an anecdote about some dangerous behavior, such as throwing rocks,

that led to a bad outcome for the person who engaged in the behavior.

John Malouff and Sally Rooke

See also Observational Learning; Operant Conditioning; Reinforcement; Social Learning Theory

Further Readings

Bandura, A. (1986). *Social foundations of thought and action.* Englewood Cliffs, NJ: Prentice Hall.

Hayes, S. C. (Ed.). (2004). *Rule-governed behavior: Cognition, contingencies, and instructional control.* Reno, NV: Context.

Ollendick, T. H., Daily, D., & Shapiro, E. S. (1983). Vicarious reinforcement: Expected and unexpected effects. *Journal of Applied Behavior Analysis, 16*, 485–491.

Powell, R. A., Symbaluk, D. G., & MacDonald, S. E. (2002). *Introduction to learning and behavior.* Belmont, CA: Wadsworth.

Rooke, S. E., & Malouff, J. M. (2006). The efficacy of symbolic modeling and vicarious reinforcement in increasing coping-method adherence. *Behavior Therapy, 37*, 406–415.

VIRTUAL SCHOOLS

Few teachers, parents, or students would disagree with the observation that technology has had a profound impact on the education of individuals at all levels. Distance learning, where students and their instructors can be separated by thousands of miles, is only one example. Virtual schools are another.

A *virtual school* is a school that provides identical content to the traditional school building but is not constrained by a physical plant or structure. Virtual schools operate electronically, providing curricula, opportunities for parent involvement, and other essential services to individuals (mostly school-age children) at home because they are being homeschooled or at home because they cannot attend school for medical or legal reasons.

In effect, virtual schools have opened up the opportunity to millions of students to pursue public school educations in a way that combines new technology and appropriate pedagogy and content. According to a recent article in *Education Week*, these are "anywhere, anytime" schools.

Currently, more than a million students are enrolled in a virtual schools in the United States, and given the increasing interest in them from families

homeschooling their children and from others, the number is expected to increase in coming years. Such students are particularly attractive to school districts (who often get funded based on their enrollment figures) because students who enroll in a virtual school are often those who would be unaccounted for if that possibility did not exist. For example, parents with strong religious ties who would rather homeschool their child can easily do so through a virtual school, saving them enormous commitments of time and expense. Educating children through virtual schools reduces planning time and cuts down on expenses associated with educating children (supplies, computers, infrastructure, etc.).

Virtual schools tend to operate as follows. Students enroll in kindergarten through 12th grade, as they would in nonvirtual schools, and parents and the student as a group are invited to an orientation (as would be case with traditionally schooled students). Books and usually computers and printers are given out upon enrollment, and a teacher is assigned who will deliver assignments and play an active role in the teaching process through initiating and monitoring assignments.

The mechanisms by which virtual schools are funded varies as greatly as the design of the school. A combination of local and state funds, tuition costs to parents, and federal funds (which almost every school receives) make up the bulk of the support.

Do virtual schools "work"? If the definition of *work* is the completion of a curriculum that otherwise would not have been completed, the answer is probably a resounding yes. However, the movement is still too new such that there are not sufficient data to conclude that large amounts of eligible students would enroll when they would not have otherwise if the option were not available.

However, there are other measures of success for a school, which also need to be examined. Little is known about the impact that virtual schools have on academic achievement, participation in extracurricular activities, and the benefits (and disadvantages) of socializing throughout the day with peers.

Neil J. Salkind

See also Distance Learning; Learning; Learning Communities

Further Readings

Barbour, M. K. (2006). [Review of the book *Development and management of virtual schools: Issues and trends*].

The International Review of Research in Open and Distance Learning, 7(1). Retrieved June 18, 2007, from http://www.irrodl.org/index.php/irrodl/article/viewArticle/307/482

Ehrich, R. W., & Kavanaugh, A. L. (2000). *Managing the evolution of a virtual school.* Norwood, MA: Artech House.

Harden, R. M., & Hart, I. R. (2002). An international virtual medical school (IVIMEDS): The future for medical education? *Medical Teacher*, 24(3), 261–267.

VOCATIONAL EDUCATION

As commonly employed, the term *vocational education* can suggest a range of diverse connotations such as "industrial education" or "education for the academically less able." However, for present purposes, this entry will use, provisionally, the mainstream sense of *vocational education*, that of "a formal course that equips students to engage in a specific occupation, trade, or profession." Following this definition, the most common vocational education courses are obviously those that prepare novices for entry to an occupation, trade, or profession. These are supplemented by still other vocational education courses that aim to increase the levels of knowledge, skill, and capability of existing practitioners, so as to enable them to carry out their occupation, trade, or profession at higher levels of proficiency.

According to this characterization, vocational education is vital for the production of a highly skilled and proficient workforce. Given the complex structure of labor markets in modern societies, and the enormous range of skills and abilities called for by labor markets, vocational education ought to be a crucial centerpiece of overall education systems. Unfortunately, throughout most of the history of formal education, and particularly since schooling became compulsory, the role of vocational education vis-à-vis the mainstream of formal education has been very problematic. Vocational education has rarely been a centerpiece of the edifice of formal education.

However, it should also be noted that the given definition of *vocational education*, like any definition, reflects various cultural and political assumptions. Its focus on vocational education as essentially the means of developing knowledge, skills, and capabilities required for occupations, trades, or professions is very consistent with contemporary, neoliberal economic agendas. But this is a partial approach in that it views vocational education in purely economic terms, thereby omitting the interests of the worker and the wider social significance of work. Yet, during much of the history of vocational education, the idea of work as a vocation or calling has been crucial. This has included considerations about the wider role of the worker as a citizen. A common theme has been that workers of all kinds should gain personal growth and satisfaction from their participation both in work and in their community. This idea links to the traditional notion of a vocation as a calling. It suggests that work needs to be rewarding in more senses than mere monetary ones. This dimension of vocational education accords with the richness implied by the term *education* as opposed to *training*. To gain a fuller understanding of vocational education and the forces that have shaped it, it is important to consider some history.

History

If educational processes are thought of broadly as the upbringing and instruction of the young in the basic skills of living, survival, and reproduction, a process that is needed to secure the continuity of any society into the next generation, then vocational education is probably older than any other kind of education. The nomadic existence of primitive societies centered on such activities as hunting and food gathering within distinct groups or communities. So early educational processes were no doubt focused on passing on to the young people of these communities the elements of such key activities as hunting, fishing, preparing food, making clothing, and caring for the young. Later on, as the development of agriculture enabled people to live in one place more or less permanently, villages and then towns were created. This led to educational processes becoming more complex as job specialization intensified. The more generic roles involved in hunting and food gathering gave way to a diversity of more specialist roles. The various basic trades developed, and with them came the concept of apprenticeship. The institution of formal and legal arrangements for apprenticeships can be regarded as the beginnings of formal vocational education and training (VET). Although the learning continued to be "on the job," the formal specification of the range of learning that the master was expected to provide for the apprentice represented an important step beyond the mere informal passing on of manual skills from one generation

to the next. Alongside of these emerging VET arrangements were the beginnings of what was to become university (or higher) education. This centered on religious institutions as major centers of learning, with a particular focus on the preparation of entrants to a priestly or religious vocation. Later, the other traditional professions, particularly law and medicine, became part of these early higher education arrangements. However, the rapid growth of what was to become the formal education system awaited the stimulus of the industrial revolution. With the rapid creation of increasing numbers of jobs that required significant literacy and numeracy, first elementary (or primary) schooling became more widespread and, eventually, compulsory. This pattern was repeated later by secondary schooling. Yet, over the lengthy time during which the formal educational mainstream (primary, secondary, higher education) developed, in most cases VET remained curiously isolated from this mainstream. This is despite the fact that VET was the means whereby the vast majority of people came to learn their occupational skills. Apprenticeships and other VET courses had become significantly more formal, though many, perhaps most, occupations were still learned informally, on the job.

Two especially prominent conceptual dualisms served to shape the thinking behind the development of the formal education mainstream, as well as the widespread exclusion of vocational education from this mainstream. These dualisms are mind versus body and theory versus practice. Each of these and their ongoing influence on VET will now be considered in some detail.

Mind–Body Dualism

Much educational theory and practice has been strongly shaped by the view that the development of minds is the main function of education. This assumption goes back to the Greeks, particularly Plato and Aristotle, who elevated theoretical knowledge and understanding over both practical and productive knowledge and know-how. The later, strong influence of Cartesian ideas on educational thought only served to reinforce this assumption. Humans are viewed as essentially minds that only incidentally inhabit material bodies. Thus, the development of mind became the main focus of education. Also, for Cartesians, thinking, as the essential feature of minds, can be divorced from other, nonessential features such as

habits, emotions, and the will. As a result, propositional learning (mindful learning) is regarded as more educational than skill learning (bodily learning). Thus, propositional learning is held to be true education, whereas skill learning is viewed as mere training; according to this view, the role of thinking within skilled performance is minimal. Presumably few people would really want, say, their home to be constructed by people who employ only physical skills with no thinking involved, yet education policy and practice continue to exemplify the ongoing influence of this mind–body separation.

A further important outcome of mindful learning being valued over bodily learning was that a class division of occupations was created. Work that was perceived to be more mindful was accorded higher status and attracted more rewarding working conditions, whereas work that was perceived to be more physical or manual in nature was accorded lower status and given less rewarding working conditions. As part of this class divide, the term *vocational education* generally came to be applied only to the preparation of entrants for those occupations perceived to be of lower status. The term *higher education* was increasingly applied to entry arrangements for professional, technological, and other high-status occupations. Thus, vocational education too often came to be viewed as the less desirable alternative, an option for those unable to gain entry to higher education. This in turn reinforced the previously noted trend to debase vocational education by portraying it as mere training rather than as genuine education. Over time, more and more emerging technical and administrative occupations have sought to demonstrate that they require high levels of mindful learning, thereby seeking to have their entry education arrangements covered by higher education provisions. This has generated a paradox in relation to occupational earning power. If the level of earning power is taken to indicate vocational success, then higher education courses are more vocational than are vocational education courses, as graduates of higher education almost invariably earn higher remuneration.

Thus, the earlier provisional definition of *vocational education* as a formal course that equips students to engage in a specific occupation, trade, or profession clearly needs amendment. It covers, in most instances, a restricted range of occupations, mostly ones that are thought to require a restricted range of propositional understanding and, hence, are of relatively low status.

It is true that much recent educational policy has sought to promote clearer pathways and links between higher education and vocational education. However, the traditional institutional segregation is still largely intact and is reinforced by the conceptual dualisms, such as mind versus body and theory versus practice that have dominated educational thought. Thus, dualisms (that are themselves merely conceptual) shape real-world practices and institutions, thereby becoming concrete dichotomies. Although strongly attacked by influential writers such as Dewey, the impact of these dualisms on educational thought and policy has remained stubbornly persistent.

Impacts on Understanding of Vocational Education

Behaviorism

Behaviorists reacted against "introspective psychology" by attempting to provide a scientific study of human behavior. They focused on observable and measurable operant behavior, on its relation to stimuli that precede the behavior, and on consequences of the behavior. This focus on observable and measurable operationally defined behavior has the seductive advantages of reliability of measurements and validation of the efficacy of procedures via systematic replication. However, because of the varying class-based influences of mind–body considerations on the education of entrants to occupations, the influence of behaviorism has been somewhat mixed. Those lower-status occupations that generally fall under the term *vocational education* and emphasize skill (or bodily) learning have typically been touched significantly by behaviorism. Whereas for those higher-status occupations that emphasize mindful learning and therefore tend to be located within higher education, the influence of behaviorism has been noticeably less. For the former, attempts to treat all of the vocational learning in behavioral terms have been common. For the latter, the influence of behaviorism has been generally confined to the more manual aspects of the occupation, for example, hand-washing routines in health-related occupations.

However, even in those vocational education courses that emphasize skill (or bodily) learning, behavioristic approaches to teaching and curriculum have had limited lasting success. Such success seems to be restricted to repetitive tasks that can be readily routinized and are relatively context-invariant; for example, performance of basic operations on a photocopier.

But most work involves much more than discrete, repetitive, routine tasks. These aspects of work that go beyond the routine require workers to display higher-order integrative capacities that involve reading changing situations and adjusting or adapting accordingly. It is the application of these kinds of capacities that does not sit well with behaviorism. Although current rhetoric about worker adaptability and creativity may be exaggerated for many occupations, there is enough truth in it to challenge behaviorism.

The worldwide trend, over the past decade, toward competency-based training (CBT) has served further to emphasize the limitations of behaviorism. CBT aims to describe precisely the elements and functions that together constitute proficient practice within an occupation. This competency description is then used as the basis for designing the entry-level training course. Thus, CBT is primarily a mechanism that attempts to ensure that vocational courses directly meet the supposed needs of industry. But CBT has not been restricted solely to the occupations that are encompassed by vocational education. It has been adopted also in various parts of the world by professions, thereby affecting higher education as well. With its focus on competence rather than behavior as such, the language that CBT uses to describe work elements and functions tends to center on human attributes such as skills, abilities, and capacities. However, it is not difficult to identify versions of CBT that are directly descended from behaviorism. Broadly speaking, there have been three distinct models for implementing CBT: task, generic, and integrated.

Task Model. This model describes work by way of listing all of the tasks performed in the occupation. The result is usually a lengthy compilation of discrete, and often minute, physical tasks. Because performance of each of these discrete tasks is directly observable and assessable, the links with behaviorism are clear. However, experience has shown that this approach to CBT has severe limitations. Capacity to perform each of the many discrete tasks does not add up to occupational competence. Occupational competence is about the capacity to seamlessly integrate the discrete competencies into a more holistic performance that matches the requirements of a given set of circumstances. The task-based model commonly confuses performance of tasks (which is observable) with

possession of skills, abilities, and capacities, which is not directly observable at all; rather, possession of skills is inferred on the basis of observed performance. In addition, the exact nature of the skills, abilities, and capacities needed for competence in a given occupation are somewhat contested. Thus, the supposed objectivity of this model is somewhat illusory. Despite matching peoples' initial commonsense intuitions about CBT, this model has proved to be a poor basis for developing such a program.

Generic Model. This model seeks to overcome the problem of innumerable discrete competencies, which plagued the task model, by identifying the generic competencies (or skills) that are crucial to competent performance of a particular occupation. It achieves the desired reduction in the number of competencies, but with the absurd consequence that very any occupations start to look almost indistinguishable. After all, many jobs require workers to communicate; to gather, organize, and analyze information; to problem solve; to use technology; to work with others; and so on. So the generic model is in danger of being too general for either training or assessment purposes. There are also the twin problems that the generic competencies only gain meaning when located in particular contexts, but when this is done, they form holistic clusters rather than being distinct competencies. So, typically, a worker might be simultaneously communicating, problem solving, and gathering and organizing information all within a particular context. Performance on isolated generic competencies is not necessarily a good indicator of capacity to perform in these more complex real-life situations. Thus, the generic model of CBT has proved to have distinct limitations. However, its reach has extended to the higher education sector where numerous universities have thought it politic to identify and advertise the generic competencies (or "graduate attributes") claimed to be common to all of their graduates. This has occurred despite the weaknesses of the generic model.

Integrated Model. This model focuses on the substantial key tasks or functions that are involved in an occupation and combines these with an account of the main skills, abilities, and capacities that are required to perform these key tasks. Thus, this model integrates selected tasks with the personal attributes needed by a competent practitioner. It represents a more holistic, phenomenological approach to CBT. The most successful examples of CBT practice tend to reflect this model.

Overall, despite the continuing success of applied behavioral analysis with particular groups, such as adults with intellectual disabilities, the continuing impact of behaviorism on vocational education is minor.

Cognitive Approaches

Cognitive approaches provide a further example of the impact of mind–body dualism on understandings of vocational education. Reacting against the behaviorist rejection of the mind, early cognitive scientists placed the mind at the center of understandings of vocational performance. Influenced by the model of the mind as a computer or information processor, these early cognitivists sought to understand mental events as a basis for explaining action or performance at work. Thus, cognitive approaches focus on individual minds and their inputs and outputs.

However, advances in both cognitive science and other disciplines led to this somewhat simple picture becoming much more complex. Recent work in cognitive science shows that much learning is unconscious—not just bodily learning of the kind favored by behaviorists but also cognitive activities such as pattern recognition. Likewise, research on expertise reveals that much of it is tacit; for example, highly skilled performers "know more than they can say." In these ways, cognitive approaches to understanding vocational performance have become more multidisciplinary and eclectic. Concepts such as meaning making and judgment have become prominent in recent work.

The implausibility of mind–body dualism has become more and more apparent with advances in neuroscience. Human beings do not consist of two disparate parts, a mind and a body, that somehow need to be connected. Whole persons perform vocational roles. Whereas reasoning, for example, is a function of a person's mind or brain, it does not follow that the person's reasoning activities are explicable solely in terms of the brain.

Theory–Practice Dualism

The theory–practice dualism is a co-relative of the assumption that mindful, or propositional, learning is the best kind of learning. This dualism provides the popular understanding of high-level practice; that is, practice that appears to involve significant cognitive content, such as the practice of professionals. Quite simply, the theory–practice dualism explains relatively

complex workplace performance by conceptualizing such demanding practice as application of theory. The sovereignty of theory over practice is thereby preserved. This view, which seeks to reduce practice to theory, is sometimes called *technical rationality*. Technical rationality is captured in the following four propositions:

1. Practical problems have general solutions.

2. These solutions can be developed and formulated by researchers or other experts outside of the practical situations.

3. The solutions can be transmitted to practitioners, for example, in written form or via training.

4. Practitioners act to solve workplace problems by applying the general solutions.

Impact on Understanding of Vocational Education

The main impact of the theory–practice dualism on vocational education has been to entrench the front-end model as the most common means of vocational preparation. The term *front-end* refers to entry to an occupation being based on an initial period of formal education, training, or both, that needs to be completed by all novices in order to become qualified workers. Front-end courses normally occur in classrooms and practice workshops rather than in real workplaces. As *front-end* suggests, it has commonly been assumed in this model that all of the learning that is needed for a lifetime of practice can be covered in the entry-level course. The front-end model has dominated vocational preparation of all kinds, especially in professional and subprofessional occupations. Even in trades and other skilled occupations that recognize the importance of on-the-job learning for novices, the front-end model has been influential. Here, a mandatory period of formal education or training in parallel with initial work experience has become the norm.

In the recent past, the front-end model has received much criticism. Three main kinds of limitation have become evident. First, cases of incompetent practitioners, who nevertheless managed to complete the front-end course, have increasingly come to public attention. This trend has been exacerbated by a growing public willingness to challenge professionals, to expose incompetence, and to seek legal redress for unsatisfactory outcomes. Second, the front-end model

has been found wanting in occupations where what constitutes competent practice has been subject to rapid change. Too many cases have surfaced of formally certified practitioners whose practice is now hopelessly out of date. Third, a range of theoretical developments appear to clash with the front-end model and, more generally, the theory–practice dualism that underpins it.

Although the theory–practice account of work performance still retains some influence, it has been questioned increasingly, from Gilbert Ryle onward. This questioning has been sharpened by various theoretical developments that include the following:

Research on Expertise. Expert performance is more than mere application of codified knowledge; it involves learning from experience. According to the on expertise research, not all valuable knowledge in the form of propositions can be learned in advance of practice. Much of this knowledge is tacit and dispositional.

Research on Transfer. With increasingly sophisticated research designs, transfer in the usual sense turns out to be surprisingly rare. This fact has elicited proposals to replace the notion of transfer with concepts such as "'preparation for future learning." So "applying existing knowledge" is replaced by "developing new knowledge." The concept of the "knowledge society" captures the claim that successful performance of various kinds of work requires the development of new knowledge on the job.

The Practice Turn in Social and Behavioral Sciences. This "turn," which has been prominent across a range of disciplines, puts a focus on human practices as a basis for theorizing. By insisting that practice is not reducible to theory or to mental items generally, this turn specifically rejects the theory–practice account of work performance.

The emphasis on the importance of knowledge within work performance has generated a burgeoning interest in workplace learning. Of course, none of this means that formal and structured learning in classrooms and training settings has no relevance for work performance, but it is does point to the previous neglect of an important source of vocational learning.

Understanding Learning at Work

Some of the most influential early theorizing of workplace learning dates from 1970s work in organizational

psychology and management theory. Notable contributors were Chris Argyris, Donald Schön, Victoria Marsick, and Karen Watkins. Some influential distinctions and concepts arose from this work, including the following:

- The distinction between single-loop learning (where the learner exhibits reactive behavior to adapt to changing circumstances) and double-loop learning (where the learner reflectively amends or adds to previous learning in selecting a suitable course of action to deal with a challenging situation)
- The distinction between a practitioner's theory-in-use (inferred from what the practitioner actually does in given circumstances) and his or her espoused theory (the theory that the practitioner claims that his or her actions exemplify)
- The concepts of the reflective practitioner, knowing-in-action and reflecting-in-action
- The distinction between informal and incidental learning

There are some common features of this early theorizing of workplace learning, including a focus on individual learners; a focus mainly on the rational, cognitive aspects of work performance; and a tendency to view work performance as thinking or reflection followed by application of the thinking or reflection. In addition, the concept of learning itself is taken for granted rather than being theorized or problematized, resulting in workplace learning being assumed to be akin to formal learning, particularly to the acquisition of propositions. These theories downplay the role of social, organizational, and cultural factors in workplace learning and performance. These are catered to in the limited sense that they are part of the environment in which the individual learns. Later accounts of workplace learning assign greater roles to social, organizational, and cultural factors.

More recent workplace learning theories recognize that workplace learning and performance are embodied phenomena, phenomena that are significantly shaped by social, organizational, and cultural factors, factors that go well beyond the individual. For these theories, workplace learning seamlessly integrates various human qualities that encompass much more than just rationality. Thus, these theories also tend to problematize, and even rethink, traditional understandings of learning. For instance, these theories can incorporate both individual and group learning as important aspects of learning at work. Earlier writers such as Dewey and Vygotsky

have been significant influences on this recent work. Well-known examples of these more recent workplace learning theories include situated learning theories and activity theory. Situated learning theories and activity theory have been influential on much recent writing on workplace learning, writing that, though seeking new understandings, reflects the key idea that learning is socially and culturally situated.

Situated Learning Theories

The work of Jean Lave and Étienne Wenger have made commonplace the concept of workplaces as "communities of practice" and "legitimate peripheral participation" as the social learning process that novices undergo to become fully accepted members of their community of practice. Lave and Wenger have emphasized the importance of apprenticeship in learning and learning as participation in a community of practice. Rather than thinking of learning as the acquisition of discrete items, they stress learning as the novice gradually coming to function appropriately in the particular social, cultural, and physical environment of their workplace. For Lave and Wenger, learning is something that is outside of the learner's head, even outside of the body. Rather, it is located in the framework of participation; that is, in a network of relations. Lave and Wenger have been criticized for the vagueness of their key notion of community, especially as they offer it as a general account of learning. In response to this criticism, Wenger has provided a more exact account of what constitutes a community of practice. In doing so, however, he has greatly reduced the likelihood of such communities occurring in practice. This casts severe doubt on Lave and Wenger's original claim to have developed a general sociocultural account of learning. There is also doubt about how well the Lave and Wenger key metaphor of participation accounts for change. As the case of various sects and religious orders shows, it is quite possible to have highly successful participation while at the same time resolutely resisting all change. It seems that Lave and Wenger have relied too much on a single factor, participation, to do the bulk of the explanatory work in their theorizing of workplace learning.

Activity Theory

Activity theorists, such as Yrjö Engeström, view workplaces as activity systems. Activity systems consist

of diverse components, such as rules, the division of labor in the workplace, and mediating artifacts. Activity systems also encompass a range of social, organizational, and cultural factors. For Engeström, learning occurs in activity systems because they continually throw up contradictions and tensions that need resolution. Thus, learning is an integral part of work processes within activity systems. Critics have questioned whether it is plausible to view all learning at work as resulting from contradictions and tensions within the activity system. Engeström's theory also posits a dialectical interplay between the learner and the activity system. But this makes unclear the extent to which the learner is the locus of learning as against the system being the locus. In his later work on activity systems, Engeström identifies a trend showing that collaborative expertise is becoming increasingly important in workplaces, whereas individual expertise is becoming less important. Such collaborative expertise is said to be facilitated by workers being prepared to cross established boundaries and negotiate with diverse other workers to improvise collaborations.

Conclusion

What might a richer notion of vocational education look like? Certainly, it would be holistic in understanding vocational knowledge to include not just propositional understanding but also cognitive, conative, and affective capacities as well as other abilities and learned capacities, such as bodily know-how and skills of all kinds. These are all plausible components of rich work performance that involves successful action in and on the world. The learning required for such holistic performance is an ongoing process rather than a knowledge acquisition event completed at the start of a career. This point lends some significance to the idea of lifelong learning.

Paul John Hager

See also Adult Learning; Cognitive View of Learning; Learning; Personality Tests; Social Learning Theory

Further Readings

Arguelles, A. (2000). *Competency based education and training: A world perspective.* Mexico City: Noriega/Limusa.

Argyris, C., & Schön, D. A. (1978). *Organizational learning: A theory of action perspective.* Reading, MA: Addison-Wesley.

Beckett, D., & Hager, P. (2002). *Life, work and learning: Practice in postmodernity.* London: Routledge.

Engeström, Y. (1999). Activity theory and individual and social transformation. In Y. Engeström, R. Miettinen, & R. Punamaki (Eds.), *Perspectives on activity theory.* Cambridge, UK: Cambridge University Press.

Hager, P. (2004). Front-loading, workplace learning and skill development. *Educational Philosophy and Theory*, *36*(5), 523–534.

Hager, P. (2005). Current theories of workplace learning: A critical assessment. In N. Bascia, A. Cumming, A. Datnow, K. Leithwood, & D. Livingstone (Eds.), *International handbook of educational policy* (Part 2, pp. 829–846). Dordrecht, Netherlands: Springer.

Lave, J., & Wenger, E. (1991). *Situated learning.* Cambridge, UK: Cambridge University Press.

Marsick, V., & Watkins, K. (1990). *Informal and incidental learning in the workplace.* London: Routledge.

Ryle, G. (1963). *The concept of mind.* Harmondsworth, UK: Penguin. (Original work published 1949)

Stevenson, J. (Ed.). (2003). *Developing vocational expertise.* Crows Nest, NSW, Australia: Allen & Unwin.

VOUCHERS

Vouchers (also referred to as *educational* or *school vouchers*) are a policy and practice by which financial credits (sometimes called *scholarships*) are awarded to students so that they can select the school that they wish to attend and use the voucher to pay for tuition and other costs. Vouchers are a direct growth out of a free-market philosophy characterized by the Nobel laureate Milton Freidman in his book *Free to Choose* and are a political mainstay of several conservative movements during the past 20 years. The idea is based on the philosophy that if parents are unsatisfied, then they can "shop around" to select the school they think is best for their children. School vouchers have been important as the basis for the creation of charter and magnet schools as well.

According to several sources, including the popular blog citizenJoe, there are at least six states, plus the District of Columbia, that currently have voucher programs, with Utah, Ohio, and Wisconsin being the most recent to add such a program. Popular in the early 1990s, voucher programs have lost some of their popularity, with relatively few states offering new programs in recent years.

Most often, school vouchers are intended for children from families who are low income, as defined by

federal poverty standards, or for children who attend failing schools. Each child is given a varying amount of tuition credits ranging from $2,700 in Cleveland up to $7,500 in the District of Columbia. It should be noted that there have been significant legal and political challenges to the implementation of school vouchers and that the issues remain unsettled.

Concerns by those opposed to the use of vouchers have to do with the use of public school money to finance selective or exclusionary programs and the potential for public funding to be applied to parochial or faith-based educational programs.

However, school vouchers have received an increasingly large amount of attention and funding. Toward the end of the 1990s, vouchers were being awarded to more than 100,000 students from a fund of more than $300 million.

A significant concern is how children who have special needs will be treated through the incorporation of such programs. In an extensive review of the impact of vouchers on children with disabilities by the National Council on Disability, a final report concluded that several factors in the acceptance of such programs in the public schools should be considered.

- In general, vouchers will not extend to children and youth with disabilities. Relevant parts of the Rehabilitation Act and the Americans with Disabilities Act will still apply but not to most activities of the private school.
- Because only a part of the costs associated with special education are covered through vouchers, parents of children with disabilities will have to cover some additional and necessary costs. Because the majority of voucher participants are from low income settings, the probability of these children receiving all necessary services is low.
- Schools where choice is the predominant model need to follow all Americans with Disabilities Act statutes; compliance will increase the costs of such programs.

Given the recent implementation of school vouchers and the lack of large-scale studies, there are few data on the effectiveness of the vouchers. For the most part, results have shown that vouchers do not have as great an impact on academic achievement as do traditional school structures. In addition, the results are mixed, not because of the actual numbers involved but because of the interpretations of the numbers. It appears that there are so many competing interests, especially where the programs are faith-based, that it will be a few years until the picture regarding effectiveness is much clearer.

Neil J. Salkind

See also Assessment; Family Influences; Parenting Styles

Further Readings

Doyle, D. P., & Finn, C. E., Jr. (1983). *Educational quality and family choice: Toward a statewide public school voucher plan.* Washington, DC: National Institute of Education.

Harmer, D. (1994). *School choice: Why we need it, how we get it.* Washington, DC: Cato Institute.

McGroarty, D. (1996). *Break these chains: The battle for school choice.* Rocklin, CA: Prima.

Skillen, J. W. (Ed.). (1993). *The school-choice controversy: What is constitutional?* Washington, DC: Center for Public Justice.

Web Sites

citizenJoe: http://www.citizenjoe.org

VYGOTSKY'S CULTURAL-HISTORICAL THEORY OF DEVELOPMENT

Lev Semenovich Vygotsky (1896–1934) was a literary scholar turned psychologist. He was an integrative thinker who conducted research and analyzed theoretical issues during a brief postrevolutionary career in Russia (1924–1934). Vygotsky focused on understanding the development of higher forms of cognition as those processes are influenced by culture—specifically, the signs and symbols of one's culture. In his 10-year career, Vygotsky wrote on the cognitive difficulties of disabled children, developed a single-subject research method, analyzed thinking and speech, and developed his theory of cognitive development, which he designated as *cultural-historical*. He did not refer to his theory as *sociocultural* because that term does not reflect the child's developmental history.

Vygotsky addressed new questions for psychology that also are relevant in contemporary society. His goal was to understand the whole of human consciousness, including thinking, feeling, will, and an understanding of oneself. In his perspective, the signs and symbols of a culture, such as speech, concepts in academic

subjects, numerical systems, and advanced mathematical constructs, are instrumental in cognitive development. Vygotsky's contributions include a description of the stages, from early childhood to adolescence, in learning to use signs and symbols to master one's thinking; the importance of speech and the stages of speech in relation to thinking; the development of thinking in concepts and the pivotal role of conceptual thinking in developing higher forms of cognition; and the role of adults and teachers in fostering the child's cognitive development.

Important to Vygotsky's work was his extensive reading in several fields, including literature, psychology, philosophy, and ethnography. Reviewing the anthropological literature on early societies, Vygotsky concluded that primitive or elementary processes (involuntary attention, simple perception, and natural memory) are universal across cultures. However, higher forms of thinking vary, depending on the available symbol systems in the culture and the culture's form of reasoning with the symbols. Other early influences on his thinking were the philosopher Benedict Spinoza, who believed that rational thinking could conquer unwelcome passions, and the philosopher G. W. F. Hegel, who maintained that reality is not static but rather is always in flux. Vygotsky's writings reflect these beliefs in his statements about the essential role of conceptual thinking in understanding the world and oneself and in his description of cognitive development as constantly undergoing change. A third key influence was the view of Alexander Potebnya, a philologist, that language is a tool of thinking.

Although Vygotsky is widely cited in current publications in the United States, the major principles of his work are largely unknown. They are his single-subject research method, the principles of his theory, and his view of education.

Research Method

A major theme in Vygotsky's work was research methodology, which he maintained was essential to developing an objective understanding of human cognition. An important initial step in research, in his view, is to determine the essence or essential characteristics of the phenomenon to be studied.

Agreeing with the belief that humans had developed higher forms of thinking in the process of historical development, Vygotsky began by analyzing ethnographic writings on primitive cultures. He identified,

for example, the actions of a Kaffir man who cut notches into wood to remember the words in a missionary's sermon as the prototype of higher cognitive behavior. Through this action, the man had mastered his behavior and raised natural memory, which was inadequate for the task, to a higher level. Vygotsky concluded from this and similar examples that a higher mental structure consists of two layers. They are the stimulus-object, the task objective (e.g., the missionary's words), and the auxiliary stimulus, the stimulus-sign (e.g., notches cut into wood).

Vygotsky and his colleagues then conducted single-subject experiments on all age groups, first on memory and attention. The purpose was to determine how and when children were capable of using signs and symbols to master their thinking and the ways that these cognitive actions change throughout development. The experiments posed a difficult task, such as remembering several words on only hearing them once, with available auxiliary stimuli, such as unrelated pictures, nearby. School-age children who remembered the words by selecting a picture for each noted similarities and differences between word and picture and verbalized connections between them. For example, one school-age child chose a picture of a house to remember the word *chair* because people can sit in a house. This process of appropriating external stimuli reconstructed the child's natural memory on a higher level. The Vygotskian experiments, referred to as the *experimental-genetic method*, led to the identification of four stages that occur in developing mastery of one's thinking.

Basic Theoretical Principles

The four general stages in cognitive development, completed near the end of adolescence, transform the primitive or elementary processes of involuntary attention, simple perception, and natural memory into the highest forms of cognition. They are self-directed attention, categorical perception, thinking in concepts (conceptual thinking), and logical memory.

The first stage, in which young children unsuccessfully attempt to complete cognitive tasks directly (such as recalling a word list) is the primitive stage. In the second stage, preschool children unsuccessfully attempt to use auxiliary stimuli. However, they are unaware of the role the auxiliary stimuli should play in their thinking about the task. Stage 3, external sign use (like the notch-cutting Kaffir man) is the successful

appropriation of auxiliary stimuli to remember a word list or, in the attention experiments, to focus one's attention. Stage 4, internal sign use, is the reconstruction of memory and attention in the form of internally reconstructed cues.

Subsequent research also indicated four stages in learning to use a concept label (word) to guide one's thinking. The experiments required the subjects to identify concept examples, using the concept label (word) and a model example. The stages, which begin in early childhood and conclude in late adolescence, are syncretic images (child forms unordered heaps of objects), complexes (preschooler uses concept characteristics, such as color and shape, inconsistently), pseudoconcepts (the school-age child accurately forms groups of objects), and, finally, true conceptual thinking. Stages 3 and 4 reflect external and internal regulation of one's thinking, respectively.

Stages 3 and 4 in cognitive development reflect two levels of the mastery of one's thinking. In stage 3, the individual regulates his or her thinking externally (e.g., pictures to remember words, a model example as a guide for identifying concept examples). In contrast, stage 4 is the internal regulation of cognition, in which the individual's attention, perception, and memory are reconstructed on the basis of thinking in concepts.

The stages also reflect Vygotsky's belief that he considered three major ideas to be equivalent to each other. They are his conceptualization of higher cognitive processes, the process of cultural development, and the self-mastery of behavior by internal processes. The development of each higher cognitive process involves the reconstruction and mastery of one's thinking at a higher level, and mastery depends on the individual's appropriation of cultural signs and symbols for thinking.

Vygotsky identified three "laws" that govern the lengthy period of cognitive development. The first is the transformation from natural forms of cognition to the use of auxiliary stimuli (signs) in thinking (stage 3). The third law of development formalizes the major changes in thinking from the third to the fourth stage of sign use. The fourth stage involves the reconstruction of attention and memory on the basis of thinking in concepts and their interrelationships.

The second law describes the cultural mechanism whereby the child, and then the adolescent, undergoes stages 3 and 4 in cognitive development. From Vygotsky's perspective, cognitive development occurs on two planes. Specifically, every cognitive function (process) occurs first between two people and then within the child. In other words, every higher cognitive process was initially a relation between two people. This conceptualization differs from other perspectives that identify the internal psychological plane as the sole locus of learning and development. The two people to whom Vygotsky refers in his statement are the adult, who represents the "ideal form" of cognitive behaviors that the child is to attain, and the child, the present form of cognition. For example, in the home, the speech of parents and other caregivers represents the final or completed form. Adult speech also determines and guides the child's first efforts on the path of development.

The Role of Education

Vygotsky identified the purpose of education as developing the child's cognitive processes to a high level. Three concepts in his view of education are relevant for contemporary education. Two of Vygotsky's concepts, the zone of proximal development and the role of collaboration in the classroom, although currently attributed to Vygotsky, are frequently described in ways very different from his view. The third important aspect of Vygotsky's thinking about education is the pivotal role of subject-matter concepts in developing higher forms of cognition.

First, Vygotsky maintained that instruction can lead development when any new form of the higher cognitive processes is beginning to mature. Vygotsky referred to the diagnostic task of identifying the cognitive processes in the period of maturation as identifying the child's zones of proximal development (ZPDs). The key to identifying maturing cognitive processes is to determine the problems that the child can solve with guidance. First, the teacher selects a problem that is just beyond the child's capabilities. Then, emerging cognitive processes can be determined in any of four ways: the child's successful imitation of the teacher's solution to the problem, completion of a solution that the teacher initiates, solving the problem after the teacher explains the principles, or solving the problem with a child more advanced in mental age. In contrast to Vygotsky's definition of ZPD as diagnosing the student's maturing cognitive processes, current statements refer to the ZPD as a component of instruction, not assessment. Such descriptions also often include a variety of classroom resources—information that is not part of Vygotsky's conceptualization.

Second, current descriptions of collaboration that cite Vygotsky as the source describe it as peer collaboration in the classroom. In contrast, Vygotsky described classroom learning differently. Collaboration, in his view, is between teacher and student. The teacher serves as the "ideal form" of cognitive behavior that the student should attain. The teacher serves in the same role as the young child's caregivers earlier in his or her life. This relationship, identified in Vygotsky's second law of development, is essential for cognitive development. As Vygotsky clearly stated, the teacher's role is to work with the student on a particular question, explaining, inquiring, correcting, and then requiring the child to explain. Later, when the child works out problems in the teacher's absence, he or she independently relies on the prior exchange. However, Vygotsky did not suggest scaffolding as an instructional method. In his discussions, the practice of the teacher completing parts of a problem is a diagnostic method, not instruction.

Third, Vygotsky described the development of subject-matter concepts (which his writings refer to as "scientific" concepts) as the key to the entire history of the individual's cognitive development. A concept, such as triangle, for example, even at the simplest level, involves generalization. Unlike the child's everyday concepts, subject-matter concepts in a domain can be represented in terms of other concepts, and they form an interrelated system. Mastery of subject-matter concepts means that the student can define them easily, implement them in various logical operations, and identify the relationships among them. Triangles, for example, are the components of hexagons and other geometric figures, such as pyramids, and so on.

The importance of subject-matter concepts is two-fold. One is that science, art, and other areas can only be adequately understood through concepts. The other is that thinking in concepts reflects a high level of cognitive development and also reorganizes the other cognitive processes on a higher level. That is, the individual's self-directed attention is governed by his or her concepts, perception becomes categorical because it is influenced by conceptual thinking, and memory becomes logical because it is organized in networks of concepts.

M. E. Gredler

See also Cognitive View of Learning; Piaget's Theory of Cognitive Development; Zone of Proximal Development

Further Readings

van der Veer, R., & Valsiner, J. (1991). *Understanding Vygotsky: A quest for synthesis.* Cambridge, MA: Blackwell.

van der Veer, R., & Valsiner, J. (Eds.). (1994). *The Vygotsky reader.* Cambridge, MA: Blackwell.

Vygotsky, L. S. (1997–1999). *Collected works of L. S. Vygotsky* (Vols. 1–6). New York: Plenum.

W

Everyone has a photographic memory. Some just don't have film.

—Steven Wright

WORKING MEMORY

Working memory (or *short-term memory*) is information temporarily held accessible in the mind. It is used in the completion of mental tasks such as comprehending language, following instructions, and solving mathematical problems. Many working memory measures correlate with intelligence rather strongly, and the average capacity of working memory increases with age in childhood. An adult can concurrently hold in mind about 4 separate, simple items, or often about 7 items by using mnemonic strategies (such as remembering a telephone number by silently rehearsing it and breaking it into groups of 3 or 4 digits). Working memory is important for educational psychology in at least two ways. First, knowledge of the demands of a task on working memory helps in predicting the task difficulty. Second, knowledge of individual differences in working memory capability helps in understanding why scholastic performance varies.

Working memory differs from the vast information that one has learned over a lifetime, or long-term memory. To illustrate, suppose one knows the sentence *The quick brown fox jumped over the lazy dog.* If, on Tuesday, one recalls only that the sentence began, "The quick brown fox" whereas on Wednesday, one recalls only that the sentence ended, "jumped over the lazy dog," it is impossible to recover the message.

However, if one is able to restore the central concepts from both parts of the sentence into working memory at once, one can imagine the fox jumping over the dog. If a story problem includes too many ideas at once, the listener or reader may find it impossible to integrate them in working memory. Similarly, it would be unwise to ask a preschool child to "put the small paint brush on the middle shelf, put the large brush along with the paint on the top shelf, and move everything that was already on the middle shelf down to the bottom." One must break up this request into smaller parts to be carried out separately.

There appear to be multiple working memory mechanisms. A very small but important set of ideas can be in the focus of one's attention and awareness at once. However, working memory goes beyond what is in focus. There also are mechanisms to hold more information just beyond awareness. This may include mental representations of the progression of speech sounds in a sentence that one heard seconds ago, or the spatial arrangement of players in a basketball game one is watching. Psychologists posit temporary holding areas for such information, termed *buffers*, with different buffers for different types of information. Some researchers, such as Nelson Cowan, think of the buffers collectively as temporarily activated portions of long-term memory. Each type of mental representation may slip out of the focus of attention momentarily, but it might be recovered from the buffer. Thus, individuals may not be able to concentrate simultaneously on all

parts of a sentence they have just heard, but they might be able to repeat it by shifting their focus from one part of the sentence to the next while making use of information persisting for several seconds in a phonological buffer.

George Miller found that knowledge and understanding can help to overcome limits of working memory. This can be done by associating several items to form one larger, meaningful group or *chunk*. For example, consider a child learning the U.S. Pledge of Allegiance. Items present in working memory at once can be memorized and interassociated. At first, learning "I pledge allegiance to the flag" may heavily tax working memory. However, after the phrase is learned, it can serve as a single chunk. If the next two phrases also are learned as chunks, it becomes possible to join these three chunks using working memory: "I pledge allegiance to the flag/of the United States of America/and to the republic for which it stands . . . ." In turn, this entire sequence eventually may become one chunk, which can then be associated with further material. Working memory can be used repeatedly to build up larger and larger segments. However, to access information within a chunk, it must be unpacked. If a person is asked what letter of the alphabet comes after *f*, he or she may have to recite the alphabet from *a* to reach the right point in his or her automatic routine.

People can be abnormal in either working memory buffers or the ability to use attention. Defects in working memory for speech, and sometimes spatial information, can underlie language, reading, and mathematical disabilities. Children with various learning disabilities also often have problems staying on task or remembering what the instructions were. They may try to pay attention and yet cannot ignore distractions, personal troubles, or daydreams. Some recent research suggests that working memory task training helps children with disabilities improve educational performance.

Nelson Cowan

See also Aptitude; Attention Deficit Hyperactivity Disorder; Fluid Intelligence; Individual Differences; Learning Disabilities; Long-Term Memory; Short-Term Memory

Further Readings

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47–89). New York: Academic Press.

Conway, A. R. A., Jarrold, C., Kane, M. J., Miyake, A., & Towse, J. N. (2007). *Variation in working memory.* New York: Oxford University Press.

Cowan, N. (2005). *Working memory capacity.* Hove, East Sussex, UK: Psychology Press.

Ericsson, K. A., Chase, W. G., & Faloon, S. (1980). Acquisition of a memory skill. *Science, 208,* 1181–1182.

Kane, M. J., Brown, L. H., Little, J. C., & Silvia, P. J. (2007). For whom the mind wanders, and when: An experience-sampling study of working memory and executive control in daily life. *Psychological Science, 18,* 614–621.

Klingberg, T., Fernell, E., Olesen, P. J., Johnson, M., Gustafsson, P., Dahlström, K., et al. (2005). Computerized training of working memory in children with ADHD—A randomized, controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry, 44,* 177–186.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63,* 81–97.

Pickering, S. J. (2006). *Working memory and education.* San Diego, CA: Academic Press.

Wilhelm, O., & Engle, R. W. (Eds.). *Handbook of understanding and measuring intelligence.* London: Sage.

Z

We cannot always build the future for our youth, but we can build our youth for the future.

—Franklin Delano Roosevelt

ZONE OF PROXIMAL DEVELOPMENT

The concept of a *zone of proximal development* (ZPD) was developed by Lev Semenovich Vygotsky during the late 1920s and elaborated progressively until his death in 1934. In *Mind in Society: The Development of Higher Psychological Processes*, Vygotsky defined the ZPD as "the distance between the actual development level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peer" (p. 86). That is, the ZPD was understood by Vygotsky to describe the current or actual level of development of the child and the next level attainable through the use of mediating semiotic and environmental tools and capable adult or peer facilitation. The "size" of the ZPD was not posited as a fixed property of the child across age periods. It would, presumably, require continual assessment at each level.

An outgrowth of Zygotsky's model of psychological development, the concept of the ZPD inspired an entire generation of Russian research in the area of developmental psychology that was generally disregarded in the West, where behaviorism was in ascendance. In the latter half of the 20th century, Vygotsky's developmental model and the centrality of the ZPD in the context of teaching and learning benefited from a wider dispersion in the West, and the publication in 1962 of *Thought and Language* and in 1978 of *Mind in Society* seemed to coincide with a greater receptivity in Western pedagogical circles to sociocultural theory and led empirical support and theoretical clarity to a new wave of educational praxis that continued unabated into the next century.

The zone of proximal development was present in *Thought and Language*, but not central to the focus of that work, which was concerned principally with psychosocial development and the processes that led children from egocentric speech and spontaneous concepts to conscious learning, the use of speech as a mediating tool, and the importance of inner thought processes as vehicles for the development of (systematic and learned) scientific concepts. The elaboration of the concept of the ZPD in chapter 6 of *Mind in Society* engendered great excitement in the academic community, and a generation of educators, to greater or lesser degrees, anchored a flotilla of pedagogies and practices to Vygotsky's concepts. This latter work clearly posited the inextricable link between learning and development and emphasized the social environment (and formal schooling) as the significant mediating factor in maturational processes.

Vygotsky's Theory of Learning and Development

In Vygotsky's view, maturation was not a passive process. Thought and speech were seen as developmentally

separate processes with, as Vygotsky observed, "different genetic roots." However, speech and thought coincide at a particular developmental juncture to produce "verbal thought." Verbal thought does not exhaust the possibilities of either thought or speech—clearly there is unvoiced thought and meaningless speech—and is reflected in the effective use of tools and in problem solving. The co-occurrence of thought and speech results in the verbalization of thought and in rational speech.

The use of tools is central to human development at all stages. Tools are generally understood to be symbolic processes that the developing child gradually masters and internalizes. Not surprisingly, the mutually generative (Vygotsky used *genetic* in this sense) interaction of thought and language may be regarded, dialectically speaking, both as mechanisms and results of developmental processes. It is not a metaphorical stretch to say that Vygotsky viewed development as marked by the gradual expansion and refinement of the tools available for the acquisition of the higher psychological functions.

Speech undergoes an evolution from prerational babbling to egocentric speech and, eventually, through a process of internalization, which contributes to the development of concepts. Concept formation is triggered by the child's interaction with the social environment as he or she attempts to solve problems. The critical moment occurs when the child becomes aware of the symbolic nature of speech, apart from its nominative function. Vygotsky believed that this occurs at about age 2 and echoed contemporaries in calling this the child's "greatest discovery." At this point, the "developmental curves" of thought and language come together.

Vygotsky believed that speech development, like the development of tools generally, proceeded through four stages. The primitive or "natural" stage corresponds to the use of preintellectual speech and preverbal thought. The second stage—"naive psychology"—grows out of the child's direct physical experience and the first use of tools to solve practical problems. At this stage, the child begins to construct subordinate clauses using *because, if, when*, and *but*, thus mastering syntactic structure without really understanding the causal, temporal, or conditional relations those words represent. In short, the child masters the syntax of speech before mastering the syntax of thought. As the child accumulates experience, he or she enters a third stage, involving the use of external signs and external operations to

solve problems. Vygotsky identified the use of simple mnemonic devices (e.g., counting on the fingers) as typical of this stage. At the fourth stage, the child develops inward. The capacity to count in one's head and the use of inner, soundless speech are characteristic of this stage. Outer and inner operations elide as the child begins to focus internalized processes on the environmental (external) problems he or she faces.

Vygotsky held that all higher-order psychological processes of thought begin originally as social processes and undergo a process of *internalization.* Thinking, or the process of concept formation, undergoes an involved transition well into adolescence, and Vygotsky was very specific in *Thought and Language* about the stages and transitions of thought from childhood to adolescence. By adolescence, the developing child is able to direct his or her mental process in conscious learning and problem-solving activities.

Vygotsky differentiated between spontaneous concepts and scientific concepts. Spontaneous concepts were those generalizations derived directly from experience without benefit of systematic instruction. Scientific concepts represent the systematic and accumulated learning of humankind usually found encoded in school curricula in the natural and social sciences.

Scientific concepts may be mediated through the social experience of school, play, and other forms of interaction that encourage thinking and imitation and through the facilitation and guidance provided by more capable peers and teachers. The time of formal schooling coincides with the growing potential of learning as it becomes intentional. Via the dialectic operation of the zone of proximal development, the child moves from empirical (spontaneous) ideas to the development of scientific concepts; the egocentric verbalism of early development leads, via the interaction of thought and speech, to the development of the higher-order psychological functions.

A key to the understanding of the ZPD is recognition of the fact that Vygotsky believed that "instruction [generally] precedes development." He noted, for instance, that instruction in writing begins well before the child has developed the higher intellectual functions necessary to fully grasp its abstract function. In fact, the child has little motivation to learn writing when it is first introduced. In all cases in the curriculum, "good" instruction runs ahead of development. Instruction occurs where emerging functions are found to be immature. Children acquire habits and skills, and

as developmental processes catch up, these become internalized as scientific concepts. Vygotsky averred that development and instruction are interconnected processes that are distinct in "rhythm." One might take the view, in Vygotskian terms, that instructional organization (curriculum) is a vast social artifice—a complex tool—designed to convey a wide array of social and scientific knowledge. The child's emerging adaptational needs and abilities spur developmental progress. Vygotsky noted that instruction in a domain may proceed with little apparent change until (in his delightful phrase), "something clicks": The child apprehends a general principle, and the developmental curve rises markedly.

Vygotsky saw instruction in the different school subjects as promoting the overall development of higher functions beyond the confines of the specific domain. The higher functions to be developed through domain-specific instruction are essentially the same and may be viewed as one, complex process. Therefore, the concept of the zone of proximal development provides an overall principle for the organization of instruction within the domains of formal schooling and with regard to whole child development through the experience of formal schooling. Vygotsky remarked in *Thought and Language* that what the child can do with assistance at first, the child will eventually do unassisted. He reached the unequivocal conclusion that "the only good kind of instruction is that which marches ahead of development and leads it; it must be aimed not so much at the ripe as at the ripening functions" (p. 188). In theoretical terms, the ZPD is the motor of psychological development; in practical terms, it is the organizing principle of instructional delivery.

Uses of the Zone of Proximal Development: Disparate Views and Applications

Scholars and educators have argued whether the zone of proximal development is a metaphor, a construct, or an analogical "space" or "place" where learning occurs. Some scholars and commentators have dismissed Vygotsky's developmental theory out of hand because they believe that his communism led him to tailor his psychological theories to the strictures of political necessity. Still others have selectively embraced his concept of the ZPD without actually analyzing, and in some cases scarcely acknowledging, the implications of sociocultural theory and developmental

psychology for educational practice. This "cherry-picking" on the part of the latter group has sometimes led to careless scholarship and equivocal practice. The outright rejection of Vygotsky by the former group is itself a socially embedded value expressed as a philosophical premise. These cavils hardly invalidate Vygotsky's model.

Vygotsky's concept of the ZPD has been labeled "metaphorical" and described as a heuristic for the resolution of what and how to teach. Although not altogether incorrect, these claims are reductive. They ignore empirical evidence generated by Vygotsky, his collaborators, and later Vygotskians and fail to recognize the centrality of socialization through schooling in the maturational process of developing higher thinking skills. The ZPD can be viewed as a construct derived from a developmental theory that yields to flexible and efficacious application in a number of educational contexts. There is empirical support for both the concept of the ZPD and Vygotsky's general model of development.

It is important to recognize that poetic metaphors lead to functionally descriptive analogies. These analogies lead to models, which in turn lead to principles, corollaries, and replicable and testable, effective practices. Vygotsky's thinking, methodologically in the scientific sense and pedagogically in the most practical sense, is well beyond the purely metaphorical.

Vygotsky might have been surprised at the popular currency of his concept of the ZPD, because he most certainly would have considered his greatest achievement to be his general model of psychological development depicting a historical-genetic view of maturational processes advancing dialectically through distinct stages driven by sociocultural forces.

Dialectics of the Zone of Proximal Development

It must be observed that the concept of the zone of proximal development resolved a critical issue for Vygotsky. Jean Piaget, much admired and closely followed by Vygotsky, had posited a progressive passage through meticulously defined levels of maturational development. However, for Piaget, the process was individual and nearly entirely independent of external circumstances. No clear mechanism or cause compels childhood development in Piaget's model. In addition, there is no clear role for learning, particularly formal schooling. For Vygotsky, the ZPD is the mechanism

of development, particularly strong during the critical years of formal schooling, which propels the child through the levels of maturational development.

According to Vygotsky, children pass through "critical periods" during which their level of actual development (i.e., the current stage of maturation) comes into tension with new or emerging conceptual demands. The child undergoes no less than a restructuring of the personality—a psychological overhaul—which results in passage to a new stage. The ZPD is a unique feature of the relationship between learning and development when the child reaches the age of formal schooling. The concept of the ZPD has tremendous explicative, prescriptive, and predictive power.

Zone of Proximal Development and Dynamic Assessment

Vygotsky mounted a most cogent critique of IQ and standardized testing, although it is important to note that he did not reject the concept of "mental age" as a construct. He also seemed to have had no problem with the use of standardized tests as measures of fully developed psychological functions as indicated by independent performance.

Vygotsky's concern was mainly with the resulting focus of instruction. He believed that standardized testing tended to focus instruction on already existing abilities that were insufficient in terms of the design of an instructional response to the developmental challenge. The current practice of dynamic assessment (DA) grew directly from these observations. Vygotsky felt that standardized testing should be used in conjunction with other forms of assessment and that the most appropriate focus of the assessment process should be on identified, emerging functions of the child. As a result, DA has refocused assessment on a wider range of developmental issues, positing assisted and emerging performance as legitimate parameters of assessment.

One of the best-known experiments in the assessment of learning aptitude was conducted by the Russian psychologist Ivanova in the 1970s. In this experiment, a child was given a set of cards with pictures of geometrical forms, shapes, and colors and asked to sort the cards on the basis of their distinctive attributes. The child was given prescribed prompts from the examiner until the task was successfully completed. Thereafter, another (similar) set of cards was given to the child for the same purpose, but

without the assistance of the examiner. The "length" of the ZPD was calculated by notation of the quality and quantity of the prompts required for the child to successfully transfer the skills acquired under assistance to independent functioning.

A number of attempts have been made to develop formal DA tools. Notably, Carol Lidz and Ruthann Jepsen developed the Application of Cognitive Functions Scale (ACFS) as a dynamic assessment procedure for use with children between 3 and 5 years of age. The ACFS assesses six tasks, composed of four core scales (Classification, Auditory Memory, Visual Memory, and Sequential Pattern Completion) and two supplemental scales (Verbal Planning, Perspective Taking).

The concept of the ZPD, as seen through the approach of DA, offers an operational view of the child's actual level of development and a measure of emerging and imminent development. Utilizing the concept of the ZPD, DA unites traditional assessment, instruction, intervention, and remediation. DA is formative and prescriptive. It is interactive and encompasses an integrated vision of the teaching–learning environment.

Instructional Applications of the Zone of Proximal Development

In practical terms, the concept of the zone of proximal development refocuses instructional and teaching strategies on maturing psychological functions. In effect, the instructional task in the light of the ZPD is to facilitate the development of the *whole child.*

Instructional design might begin with a child's actual level of development. During this process, an assessment of the child's potential for moving on to higher levels of functioning should be made. Eventually, the focus of instruction becomes the zone of proximal development and the identified targets of the next level. Finally, the child reaches a new level of development and the process begins again.

Learning, then, is a social process. Vygotsky rested his theory of cognitive development on the process of *internalization.* The child engages in a social process with a guide or more skilled peer and eventually, through *imitation*, is able to perform higher functions autonomously. The concept of imitation held, for Vygotsky, a special and technical meaning closely related to the concept of the ZPD. The child may not be expected to imitate behaviors that lie outside of his or her zone of proximal development. Those behaviors

that the child is able to intellectually imitate, particularly when assisted by a more capable guide, are indicators of emerging functions and demarcate the ZPD. Vygotsky uses the example of the initial efforts of the novice to learn chess. A child may play checkers with a chess set as a result of a conceptual generalization based on color. In developmental terms, the child has failed to acquire, at this point, a scientific concept of roles and functions indicated by the identities of the chess pieces.

In the determination of the ZPD, interaction with the child is critical. Imitation and collaboration provide the basis for tailoring the teaching-learning experience of the child to developmental needs. *Collaboration*, in the Vygotskian sense, does not simply mean joint or coordinated effort. It refers to situations in which the child has the opportunity to interact in social contexts for the purpose of problem solving. Collaboration is necessary, in this sense, because it maintains the dynamic tension between what the child is able to achieve independently and what the child may achieve through imitation and the gradual apprehension of modeled behaviors.

Vygotsky was not specific about the role of the adult guide in the interactive environment of the ZPD. He mentioned only four guiding behaviors in his writing on the subject: showing how to solve a problem and guiding the child in imitation; beginning a problem and allowing a child to finish; proposing to a child that he or she solve a problem together with another, more developed child; and explaining a principle for solving a problem, asking leading questions, and analyzing a problem. What is clear is that learning is enhanced in collaborative and guided environments like the circumstances of formal schooling. Autonomous behavior and higher-order psychological functions begin in a social environment and are internalized by the maturing learner.

The ZPD, clearly, has diagnostic and prescriptive uses, not all of which were contemplated by Vygotsky. As mentioned, many of these uses are analogical, metaphorical, or schematic and do not necessarily integrate Vygotsky's developmental model of sociocultural learning with specific practice.

Vygotsky's Methodology

Vygotsky's overriding concern was the development of the "higher psychological processes." His critique of stimulus-response frameworks, which included Russian

psychology (reactology), German behaviorism (Wilhelm Wundt), and Western behaviorism (J. B. Watson, Edward L. Thorndike), centered on the fact that these models usually focus on the etiology of lower psychological functions—those shared by humans and anthropoids. Vygotsky's theoretical focus was on the distinctly human characteristics of higher psychological functioning. His methodology rested on three theoretical dispositions: (1) a focus on the analysis of process rather than the observation and analysis of stable, fixed objects; (2) an emphasis on the explanation of essential and defining characteristics of human psychological development rather than a focus on the merely descriptive and superficially observable; and (3) a focus on the "historical" study of the process of change. Vygotsky believed that much of the psychological research of the time was brought to bear on "fossilized behaviors" that had become automatic or mechanized (i.e., autonomic). In short, he proposed a psychology that focused on the analysis of processes, the explication of the dynamics of human development, and the reconstruction and analysis of critical developmental junctures.

Experimentally, Vygotsky's model of the process of maturation led him to what he called, in *Mind in Society*, "the functional method of double stimulation." This approach, as seen in the example discussed with relation to the topic of DA, allowed Vygotsky and later researchers to identify the inner structure of current stages of maturation and to identify proximal stages.

The experimental methodology of Vygotsky and later Russian psychologists influenced by him has not been entirely translated into English. Vygotsky, particularly in his translated work, tended to report his conclusions and research methodology separately; although he uses frequent examples in the extant English language sources, they are usually unaccompanied by the observational data. Later Vygotskians broadened the experimental base of his model, although much of that body of research has yet to be translated into English.

Enduring Influence of the Zone of Proximal Development

The influence of Vygotsky's concepts on modern educational practice can be recognized in the development of new models in instructional design. These models claim Vygotsky as a forbearer to the direct application of his ideas to the challenge of instructional design and delivery in the modern classroom environment.

There is a definite relationship between Vygotsky's sociocultural theory of the ZPD and the proponents of collaborative learning, as Vygotsky provides the epistemological model required to ground collaborative strategies in a theory of learning. In much the same way, constructivists of all types and inclinations tend to draw an unbroken line from Vygotsky, Piaget, and John Dewey to more recent expositors like Jerome Bruner, and then finally to current practices, without carefully distinguishing the unique and significant contributions of each.

More directly, the approaches generally known as *situated cognition, cognitive apprenticeship*, and *scaffolding* (and related constructs of *modeling, coaching*, and *fading*) find theoretical support and epistemological grounding in the concept of the ZPD.

Vygotsky's achievement in developing the ZPD concept lies in the fact that he successfully combined a theory of child development and a psychosocial model of mind which, together, focus the process of teaching and learning directly on the child in his or her social context. In so doing, he may be regarded—along with John Pestalozzi, Maria Montessori, Dewey, and Piaget—as a progenitor of the child-centered view of modern education. With the ZPD, Vygotsky's legacy must be considered fourfold:

1. The establishment of an organizing principle for instruction that relates whole child development in a social setting to specific learner goals to complete academic tasks and master learning domains

2. The redefinition of the relationship between teachers and learners

3. The redirection and broadening of assessment criteria

4. The establishment of a critical link between formal schooling and psychological development

The concept of the ZPD has been extremely powerful in the generation of systematic methodologies and practical strategies. Notwithstanding, the idea that the human mind is powerfully shaped by social forces and the interaction of the child with his or her environment is unsettling to Western, especially American, educators who see it as contrary to half-formed, frequently nebulous and ill-defined value formations, such as independence, American individualism, and self-reliance. This accounts for the fact that many practitioners have embraced the ZPD and practices derived from the concept itself, while neglecting or rejecting Vygotsky's sociocultural point of departure and the developmental implications of the teaching–learning process.

John J. Ketterer

See also Learning; Learning Strategies; Learning Style; Piaget's Theory of Cognitive Development

Further Readings

Berk, L. E., & Winsler, A. (1995). *Scaffolding children's learning: Vygotsky and early childhood education.* Washington, DC: National Association for the Education of Young People.

Cole, M., & Wertsch, J. V. (2004). *Beyond the individual-social antinomy in discussions of Piaget and Vygotsky.* Retrieved January 21, 2007, from http://www.massey.ac.nz/~alock/virtual/colevyg.htm

Daniels, H. (2001). *Vygotsky and pedagogy.* London: Taylor & Francis.

Gredler, M., & Shields, C. (2004). Does no one read Vygotsky's words? Commentary on Glassman. *Educational Researcher*, *33*(4), 21–25.

Lidz, C. S., & Jepsen, R. H. (2000). *The Application of Cognitive Functions Scale.* Unpublished manuscript.

Marsh, G. E., II, & Ketterer, J. J. (2005, Summer). Situating the zone of proximal development. *The Online Journal of Distance Learning Administration*, 8(2). Retrieved December 12, 2007, from http://www.westga.edu/%7Edistance/ojdla/summer82/marsh82.htm

Piaget, J. (1955). *Language and thought of the child* (2nd ed.). New York: Meridian Books.

Rogoff, B., & Wertsch, J. V. (Eds.). (1984). *Children's learning in the "zone of proximal development."* San Francisco: Jossey-Bass.

Tudge, J., & Scrimsher, S. (2003). Lev S. Vygotsky on education: A cultural-historical, interpersonal, and individual approach to development. In B. J. Zimmerman & D. H. Schunk (Eds.), *Educational psychology: A century of contributions* (pp. 207–228). Mahwah, NJ: Lawrence Erlbaum.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.). Cambridge, MA: Harvard University Press.

Vygotsky, L. S. (1986). *Thought and language* (A. Kozulin, Ed. & Trans.). Cambridge: MIT Press. (Original work published 1962)

Index

Entry titles are in bold.